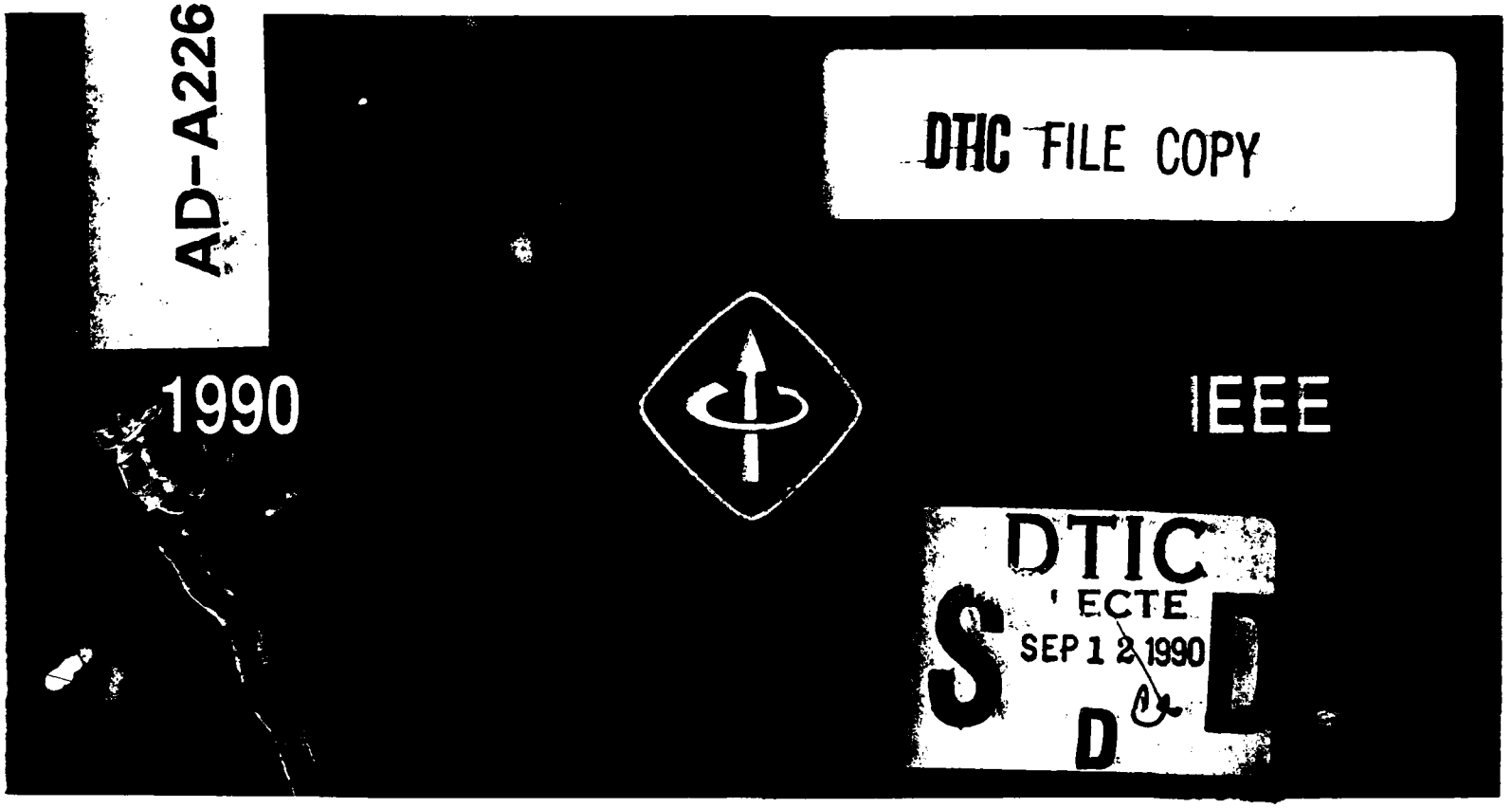
ABSTRACTS OF PAPERS

(1)



INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY

JANUARY 14-19, 1990

SHERATON HARBOR ISLAND EAST HOTEL
SAN DIEGO, CALIFORNIA

SPONSORED BY: IEEE INFORMATION THEORY SOCIETY

IEEE Catalog Number 90 CH 2711-0

1990 IEEE INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY

(ISIT)

SHERATON HARBOR ISLAND EAST HOTEL
SAN DIEGO, CALIFORNIA

JANUARY 14-19, 1990

Sponsored by:

The Institute of Electrical
and Electronics Engineers,
Information Theory Society

Co-Chairmen

L.B. Milstein A.J. Viterbi

International Advisory Committee

J.K. Wolf	(USA-Chairman)
R. Ahlswede	(West Germany)
I.F. Blake	(Canada)
G.C. Coraxxa	(Italy)
M. Costa	(Brazil)
I. Csiszár	(Hungary)
P.G. Farrell	(United Kingdom)
I. Ingemarsson	(Sweden)
J. Justesen	(Denmark)
T. Kasami	(Japan)
T. Kløve	(Norway)
E.C. van der Meulen	(Belgium)
B. Picinbono	(France)
Z. Jeong-Pan	(China)
J. Riera	(Spain)
J.P.M. Schalkwijk	(The Netherlands)
M. Tounouga	(Cameroun)
G. Ungerboeck	(Switzerland)
A. Yaglom	(USSR)
V. Zima	(Czechoslovakia)
J. Ziv	(Israel)

IEEE Catalog Number 90 CH 2711-0 Library of Congress Number 72-179437

Program Committee

C.W. Helstrom, Chairman

F. Beutler	S.A. Kassam
E. Biglieri	J.W. Modestino
R.E. Blahutxi	R.L. Pickholtz
S. Cambanis	H.V. Poor
D.J. Costello, Jr.	E.C. Posner
T. Cover	M.B. Pursley
L. Davisson	I. Rubin
A. Gersho	J.H. Shapiro
R.M. Gray	N.J.A. Sloane
B. Hajek	D.L. Snyder
L.N. Kanal	A.D. Wyner
Finance	R. Cruz (Treasurer)
Publications	E. Masry
Registration	S. Chatterjee
Publicity	R. Lugannani
Local Arrangements	R. Rao (Chairman)
	K. Bakhru
	N. Feldman

Acknowledgment

The following organizations have provided financial support for the 1990 IEEE ISIT:

The Army Research Office
The National Science Foundation
The Office of Naval Research
Others pending.

ISIT '90 SYMPOSIUM SCHEDULE

Day	Time Period	Activity
Sunday	5-8 p.m.	Registration
	7-9:30 p.m.	Reception
Monday	8-11 a.m.	Registration
	8-8:50 a.m.	Plenary Lecture
	9-10:20 a.m.	Technical Sessions (A)
	10.20 a.m.-10:40 a.m.	Coffee Break
	10:40-12 m.	Technical Sessions (A)
	2-3:20 p.m.	Technical Sessions (P)
	3:20-3:40 p.m.	Coffee Break
	3:40-5 p.m.	Technical Sessions (P)
	7:00 p.m.	Recent Results Session
Tuesday	8-11 a.m.	Registration
	8-8:50 a.m.	Shannon Lecture
	9-10:20 a.m.	Technical Sessions (A)
	10.20 a.m.-10:40 a.m.	Coffee Break
	10:40-12 m.	Technical Sessions (A)
	2-3:20 p.m.	Technical Sessions (P)
	3:20-3:40 p.m.	Coffee Break
	3:40-5 p.m.	Technical Sessions (P)
	7:00 p.m.	Recent Results Session
Wednesday	8-11 a.m.	Registration
	8-8:50 a.m.	Plenary Lecture
	9-10:20 a.m.	Technical Sessions (A)
	10.20 a.m.-10:40 a.m.	Coffee Break
	10:40-12 m.	Technical Sessions (A)
Thursday	8-11 a.m.	Registration
	8-8:50 a.m.	Plenary Lecture
	9-10:20 a.m.	Technical Sessions (A)
	10.20 a.m.-10:40 a.m.	Coffee Break
	10:40-12 m.	Technical Sessions (A)
	2-3:20 p.m.	Technical Sessions (P)
	3:20-3:40 p.m.	Coffee Break
	3:40-5 p.m.	Technical Sessions (P)
	8:00 p.m.	Banquet
Friday	8-11 a.m.	Registration
	8-8:50 a.m.	Plenary Lecture
	9-10:20 a.m.	Technical Sessions (A)
	10.20 a.m.-10:40 a.m.	Coffee Break
	10:40-12 m.	Technical Sessions (A)

MONDAY

	1	2	3	4	5	6	7
A	Spread-Spectrum Communications	Detection Theory I	Neural Networks I	Data Compression	Channel Capacity	Coding Theory I	Trellis Coding I
P	Communication Systems	Broadcast Channels	Estimation I	Quantization I	Shannon Theory I	Coding Theory II	Error-Control and Other Coding

TUESDAY

	1	2	3	4	5	6	7
A	Communication Theory I	Modulation	Optical Communica-tions	Source Coding I	Shannon Theory II	Coding Theory III	Trellis Coding II
P	Stochastic Processes	Multiple Access I	Signal Processing I	Quantization II	Cyclic Codes	Coding Theory IV	Error-Correcung Codes I

WEDNESDAY

	1	2	3	4	5	6	7
A	Communication Theory II	Multiple Access II	Signal Processing II	Source Coding II	Shannon Theory III	Coding Theory V	Trellis Coding III

THURSDAY

	1	2	3	4	5	6	7
A	Estimation II	Multiple Access III	Cryptography I	Speech Processing	Block Coding	Coding Theory VI	Viterbi Decoders
P	Detection Theory II	Communication Networks	Neural Networks II	Cryptography II	Shannon Theory IV	Coding Theory VII	Error-Correcting Codes II

FRIDAY

	1	2	3	4	5	6	7
A	(a) Magnetic Recording (b) Information Theory Applications	Pattern Recognition	Signal Processing III	Coding Theory VIII	Entropy	Coding Theory IX	Convolutional Codes

PROGRAM

TABLE OF CONTENTS

Monday, January 15, 1990

8 - 8:50 a.m. PLENARY SESSION

9 a.m. - 12 m. TECHNICAL SESSIONS

SESSION MA1 - SPREAD SPECTRUM COMMUNICATIONS

SESSION MA2 - DETECTION THEORY I

SESSION MA3 - NEURAL NETWORKS I

SESSION MA4 - DATA COMPRESSION

SESSION MA5 - CHANNEL CAPACITY

SESSION MA6 - CODING THEORY I

SESSION MA7 - TRELLIS CODING I

Monday, January 15, 1990

2 p.m. - 5 p.m. TECHNICAL SESSIONS

SESSION MP1 - COMMUNICATION SYSTEMS

SESSION MP2 - BROADCAST CHANNELS

SESSION MP3 - ESTIMATION I

SESSION MP4 - QUANTIZATION I

SESSION MP5 - SHANNON THEORY I

SESSION MP6 - CODING THEORY II

SESSION MP7 - ERROR-CONTROL AND OTHER CODING

Tuesday, January 16, 1990

8 - 8:50 a.m. SHANNON LECTURE

9 a.m. - 12 m. TECHNICAL SESSIONS

SESSION TA1 - COMMUNICATION THEORY I

SESSION TA2 - MODULATION

SESSION TA3 - OPTICAL COMMUNICATIONS

SESSION TA4 - SOURCE CODING I

SESSION TA5 - SHANNON THEORY II

SESSION TA6 - CODING THEORY III

SESSION TA7 - TRELLIS CODING II

Tuesday, January 16, 1990

2. p.m. - 5 p.m. TECHNICAL SESSIONS

SESSION TP1 - STOCHASTIC PROCESSES

SESSION TP2 - MULTIPLE ACCESS I

SESSION TP3 - SIGNAL PROCESSING I

SESSION TP4 - QUANTIZATION II

SESSION TP5 - CYCLIC CODES

SESSION TP6 - CODING THEORY IV

SESSION TP7 - ERROR-CORRECTING CODES I

SESSION WA3 - SIGNAL PROCESSING II

SESSION WA4 - SOURCE CODING II

SESSION WA5 - SHANNON THEORY III

SESSION WA6 - CODING THEORY V

SESSION WA7 - TRELLIS CODING III

SESSION ThA2 - MULTIPLE ACCESS III

SESSION ThA3 - CRYPTOGRAPHY I

SESSION ThA4 - SPEECH PROCESSING

SESSION ThA5 - BLOCK CODING

SESSION ThA6 - CODING THEORY VI

SESSION ThA7 - VITERBI DECODERS

Thursday, January 18, 1990

2 p.m. - 5 p.m. TECHNICAL SESSIONS

SESSION ThP1 - DETECTION THEORY II

SESSION ThP2 - COMMUNICATION NETWORKS

SESSION ThP3 - NEURAL NETWORKS II

SESSION ThP4 - CRYPTOGRAPHY II

SESSION ThP5 - SHANNON THEORY IV

SESSION ThP6 - CODING THEORY VII

SESSION FA2 - PATTERN RECOGNITION

SESSION FA3 - SIGNAL PROCESSING III

SESSION FA4 - CODING THEORY VIII

SESSION FA5 - ENTROPY

SESSION FA6 - CODING THEORY IX

SESSION FA7 - CONVOLUTIONAL CODES

PLENARY SESSION

Monday, 8 - 8:50 a.m.

Computational Complexity as a Scientific Metaphor

Christos H. Papadimitriou *University of California, San Diego; La Jolla, CA 92093*

Computational Complexity is a mathematical theory which strives to explain the large amounts of computational resources (typically, computation time) apparently required for the solution of many important computational problems. During the past twenty years, research in Complexity has focussed on problems of combinatorial nature, with occasional excursions into algebraic and numerical computation. One of the achievements of this theory has been the classification of computational problems according to their potential for an efficient solution. Of these classes, perhaps best known among researchers in other fields is that of *NP*-complete problems, a class of problems currently believed to be intractable, requiring exponential amounts of computation.

In a literal application of these ideas, we are faced with a computational problem suspected to be difficult, and we use the tools of Complexity Theory to formally prove it intractable, that is, highly unlikely to be solvable by algorithms with acceptable performance. This talk will describe a genre of scientific argument which can be viewed as a metaphorical application of the ideas of Complexity Theory. We are faced with a difficult mathematical problem, not of a computational nature, and we use the intractability of a derived computational problem (sometimes not of direct practical interest) as evidence for the difficulty of the original (mathematical) problem. We illustrate this point by examples from several areas, including Distributed Control, Game Theory, and Operations Research.

TECHNICAL SESSIONS

Monday, 9 a.m. - 12 m.

SESSION MA1

SPREAD SPECTRUM

A Comparison of the Performance of Two Types of Narrowband Interference Rejection Techniques in DS-Spread Spectrum Systems

Sophie Y. Dayot and Laurence B. Milstein *Department of Electrical and Computer Engineering, San Diego State University, San Diego, CA 92182, and Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093*

In this paper, a comparison is made of the performance of two narrowband interference rejection schemes when used in conjunction with a direct sequence spread spectrum receiver. The first scheme is based upon the use of an estimation filter designed by solving the Wiener-Hopf equation, and the second technique makes use of a transform domain processing structure. Probabilities of bit error are presented for both systems under the conditions of either tone or narrowband Gaussian interference, and it is shown that the former scheme is superior for tone interference. However, when the interference has a finite bandwidth, in most cases the latter technique yields better performance.

Direct Sequence Spread Spectrum with Random Signature Sequences: A Large Deviations Analysis

John S. Sadowsky and Randall K. Bahr *School of Electrical Engineering, Purdue University, West Lafayette, IN 47907, and Dept. of Electrical & Computer Eng., University of Arizona, Tucson, AZ 85721*

This paper investigates the bit error probability of a direct sequence spread spectrum multiple access communications system using *large deviations theory*. Let m denote the number of interfering signals and n denote the length of the direct sequence signature sequence. We consider the limit is as $n \to \infty$ with m fixed. The interfering signals have random phases and delays with respect to the desired signal. Let Θ denote the vector of interfering signal phases and delays, and define $P_b(\theta) = \mathrm{P}(\text{bit error} \,\|\, \Theta = \theta)$. A straightforward application of modern large deviations theory indicates that $P_b(\theta)$ vanishes exponentially as $n \to \infty$ with rate $I(\theta) > 0$. Our main theoretical contribution is to prove the limit

$$P_b = E[P_b(\Theta)] \sim C\, n^{-(3m+1)/2} \exp(-\tilde{I}\, n)$$

where $\tilde{I}$ is the worst case exponential rate, that is, $\tilde{I} = \min_\theta I(\theta)$, and $\sim$ means that the ratio tends to 1 as $n \to \infty$. An important biproduct of our analysis is that it leads to numerically efficient means to estimate P_b for finite n.

Exact Analysis of Asynchronous Frequency-Hop Spread-Spectrum Multiple-Access Networks

Kyungwhoon Cheun and Wayne E. Stark *Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109*

In this paper, we derive the exact probability of error of an Asynchronous Frequency-Hop Spread Spectrum Multiple-Access (AFHSS-MA) Network transmitting one BFSK modulated bit per hop using the theory of spherically symmetric random vectors. It is assumed that Markov hopping patterns are employed. We show that the usual approximation that the probability of error is ½ (½-approximation) whenever a hop is hit by multiple-access interference, is excessively pessimistic and grossly underestimates the multiple-access capability of AFHSS-MA networks. Also, we show that using the ½-approximation has led to misleading conclusions. For example, we show that contrary to popular belief, a system using perfect side-information to erase the hops that were hit performs much worse than systems making simple hard decisions without side-information. We show that the same techniques used to derive the probability of error can be used to derive the channel statistics of an AFHSS-MA channel when

Viterbi ratio thresholding (VRT) is employed, and present numerical results to show that the VRT offers significant improvements in performance compared to simple hard decisions at a small increase in complexity.

Quadratic Congruential Coding and its Implementation in Frequency-Hop Spread-Spectrum Communication Systems

Zoran I. Kostic and Edward L. Titlebaum *Department of Electrical Engineering, University of Rochester, Rochester, NY 14627*

The code-dependent performance of frequency-hop spread-spectrum multi-user communication systems based on linear congruence (LC) codes and a subset of Reed-Solomon Codes (Costas or Einarsson's codes) is presented for the case when three sources of interference are present: mutual interference between system users (between-users interference), interference between signals originating from the same user due to the unknown signal propagation time and multiple transmissions (single-user signal interference), and doppler. Unsatisfactory performance in the case of single-user signal interference is used as a motivation for introducing a new system which uses quadratic congruence (QC) codes. The performance of this system is evaluated. It is close to the optimal in case of between-users interference, and the best known in case of single-user signal interference. Doppler interference performance is not affected by the code choice. The symmetry of QC codes is recalled and used for the design of truncated QC codes. Fullness of truncated QC codes eliminates the receiver design problems associated with non-fullness of QC codes.

"CDMA-FDMA Hybrid Protocol" for Distributed Multihop Spread-Spectrum Packet Radio Networks

Shingo Tanaka, Akihiro Kajiwara, and Masao Nakagawa *Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi Kohokuku, Yokohamashi 223, Japan*

CDMA (code-division multiple access) packet radio network has the merit that it resists fading and jamming, but also has the demerit of non-zero cross-correlation between different signals causing the "near far problem". To keep merit of CDMA and simultaneously have characteristic of anti-near far problem, we propose the "CDMA-FDMA (frequency-division multiple access) hybrid protocol" for distributed multihop packet radio networks.

In this protocol, spread frequency band is divided to some areas. In each divided band, CDMA is used. The cross-correlation of spread spectrum signals in different band equals to zero. Therefore, if we distribute some (at most 4 or 5) frequency bands to some terminals placed close to each other, higher SIR (signal-to-interference ratio) than CDMA-only protocol is expected.

>From the operation of dividing frequency band to some areas, the cross-correlation of SS signals in same band is higher than that of CDMA, because of the narrower band width than that of CDMA. But the zero cross-correlation of FDMA has higher effect than the above problem, if the distribution of frequency bands is ideal.

In four examples of networks, we explain the result that the hybrid protocol has higher SIR than CDMA only.

Error Correcting Coding and Pseudo-Random Interleaving Scheme Against Intelligent Partial Time Jammers

Philippe R. Sadot, Marc M. Darmon, and Sami Harari *Alcatel Transmissions par Faisceaux Hertziens, Military Transmissions Department, 92301 Levallois-Perret, France, and Groupe d'Etude du Codage de Toulon, Universite de Toulon et du Var, 83 130 La Garde, France*

The usual techniques of antijamming against partial time narrow band jammers consisting of spectrum spreading and error correcting coding are not efficient in the case of most military microwave

links because of the minimum data rate to transmit and of the channel bandwidth limitation.

A new strategy of antijamming is proposed, which consists in admitting that all information is erased during the short pulses of the jammer, and in trying to reconstitute the lost information by means of an error correcting coding and interleaving scheme.

In order to cope with intelligent jammers we have defined a new pseudorandom interleaving that features both characteristics of guaranteed minimum distance between symbols of the same codework and cryptographic complexity. It uses a high complexity non-linear pseudorandom generator, and a set of three rules that insure the minimum distance property.

The coding scheme and the interleaving device will be described.

The BER performance of the coding and interleaving scheme has been derived by means of algebraic techniques and the MACSYMA software, and will be explained.

Finite Memory Recursive Solutions for the Equilibrium and Transient Analysis of G/M/1-Type Markov Processes with Application to Spread Spectrum Multiple Access Networks

Garimella Rama Murthy and Edward J. Coyle *School of Electrical Engineering, Purdue University, West Lafayette, IN 47907*

G/M/1-type Markov processes provide natural models for the activity on multiple access networks. Efficient recursive solutions for the equilibrium and transient analysis of these processes are therefore of considerable interest.

In this paper, a state space expansion criterion called *level entrance direction information* is introduced for G/M/1-type Markov processes. It is then shown that this criterion leads to a new class of recursive solutions, called *finite memory recursive solutions*, for the equilibrium probabilities. A finite memory recursive solution of order k has the form

$$\bar{\pi}_{n+k} = \bar{\pi}_n \mathbf{W}_1 + \bar{\pi}_{n+1} \mathbf{W}_2 + \cdots + \bar{\pi}_{n+k-1} \mathbf{W}_k,$$

where $\bar{\pi}_n$ is the vector of limiting probabilities of states on level n of the process and $\mathbf{W}_i$, $1 \le i \le k$ are $n \times n$ matrices.

It is also shown that, utilizing the concept of LEDI completeness, a finite memory recursive solution for the Laplace transform of the vector of state occupancy probabilities at time t can be found. Such a recursive solution has the form

$$\bar{\pi}_{n+k}(s) = \bar{\pi}_n(s)\mathbf{W}_1(s) + \bar{\pi}_{n+1}(s)\mathbf{W}_2(s) + \cdots + \bar{\pi}_{n+k-1}(s)\mathbf{W}_k(s).$$

Thus, this recursive solution provides a tractable method for the transient analysis of G/M/1-type Markov processes. The relationship between finite memory recursive solutions and matrix geometric solutions is also explored.

A new G/M/1-type Markov process model of Spread Spectrum Slotted ALOHA (SSSA) networks is developed. this model provides a finite memory recursive solution for the computation of the equilibrium and transient distribution of the number of busy users. It also allows many analytical results such as the necessary and sufficient conditions for the stability and exact closed form expression for the expected delay to be found.

Moment Methods for Estimation of Fine Time Synchronization Error in FH/MFSK Systems

L. J. Mason and E. B. Felstead *Communications Research Centre, Ottawa, Canada*

Time synchronization in frequency-hopped (FH) systems consists of two stages. The coarse stage reduces the alignment error between the transmitter and the receiver hop periods to less than one half hop period. The fine stage reduces this error to typically a few percent of a hop period. Methods are described which give low-biased, consistent estimates for the fine time synchronization error under low SNR conditions. The methods are meant primarily for synchronization in a FDMA satellite system where

matched filtering and demodulation is performed by a SAW Fourier transform device followed by an envelope detector.

Two implementations, denoted the two-tone and early-late filter methods, are analyzed. For each implementation, the time error estimate is derived using either the first and second order moments of the received envelope, or the second and fourth order moments. A unique method using three second order moments for the early-late filter implementation is included. Results are given to show the effect of SNR on the mean and standard deviation of the estimates for a range of actual time error.

SESSION MA2

DETECTION THEORY I

Asymptotic Efficiencies in Multiple-Access Channels

S. Y. Miller and S. C. Schwartz *Department of Electrical Engineering, Princeton University, Princeton, NJ 08540*

The evaluation of the performance of a central multiuser receiver is an important issue in multiuser communication. In some cases, where explicit analytical results are unavailable, a fruitful approach is to resort to the (asymptotic) vanishing noise assumption which yields closed form results. In all cases, it is interesting to compare the resulting performance for any of the users to the performance of a single user receiver in a related single user channel. The possible multiple access degradation is then expressed in terms of the *efficiency* of the multiuser receiver. In this paper asymptotic efficiencies of multiuser sequence detectors are evaluated employing an approach which relies on a large deviation result of H. Chernoff. A strong relationship between the results of this approach, applied to the Gaussian multiple access channel, and existing results for optimum and sub-optimal multiuser detectors is indicated. The examples given illustrate the possible general applicability of the Chernoff result to the calculation of multiuser asymptotic efficiencies.

Importance Sampling: A Robust Statistics Approach

Geoffrey Orsak and Behnaam Aazhang *Computer and Information Technology Institute, Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892*

The problem of estimating expectations of functions of random vectors via simulation is investigated. Monte Carlo simulations, also known as simple averaging, have been employed as a direct means of estimation. A technique known as Importance Sampling can be used to modify the simulation via weighted averaging in the hope that the estimate converges more rapidly to the expected value than standard Monte Carlo simulations. The fundamental problem in Importance Sampling is to determine the appropriate density function for the underlying random variable in the simulation. Since the unconstrained optimal solution to this problem is degenerate, suboptimal solutions have been of the form of a scaled, linearly shifted or exponentially tilted version of the original density. In this paper, we derive a constrained optimal solution to the problem of minimizing the variance of the Importance Sampling estimator. This is done by finding the distribution which is "closest" to the unconstrained optimal solution in the Ali-Silvey sense. The solution from the constraint class is shown to be the least favorable density function in terms of Bayes risk against the optimal density function. Examples of constraint classes, which include ϵ-mixture, will show that the constrained optimal solution can be made arbitrarily close to the optimal solution. Applications to estimating probability of error in communication systems will be given.

Performance of Optimal Non-Gaussian Detectors

Don H. Johnson and Geoffrey Orsak *Computer and Information Technology Institute, Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892*

The optimal procedure for detecting the presence of discrete-time signals in additive noise can be derived from the likelihood ratio test. When the noise has statistically independent, identically distributed components, the dependence of the detector's performance on signal characteristics can be related to the Kullback-Leibler (KL) distance between the distributions governing the hypotheses. Performance predictions based on the central limit theorem are shown to be poor approximations to the true performance. Performance of the optimal detector has long been known to decrease exponentially with increasing Kullback-Leibler distance. symmetric noise amplitude distributions yield a symmetric dependence on the difference between the signals' amplitudes at each time index. Small-signal (locally

optimal) detection performance is shown to depend on signal energy in proportion to the Fisher information for location. When a distance measure can be defined, performance depends on a different measure than used in the detector with one exception (the Gaussian). Large-signal performance depends on the signal waveform with the most dramatic contrast between signal sets comprised of constant-difference signals and of signals differing in only one value.

Computing Distributions from Moments Using Padé Approximants

James A. Ritcey and Hamid Amindavar *Department of Electrical Engineering, University of Washington, FT-10 Seattle, WA 98195*

The evaluation of signal detection schemes is often hampered by the inability to accurately compute tail probabilities. In fact, often only the moments of the test statistic can be computed and we seek to invert this information to obtain cumulative distribution functions. One approach is to determine the moment generating function, through a truncated power series expansion. In this paper we examine the performance of Padé approximants to approximate the m.g.f., and we carry out the inversion using the method of residues. The diagonal Padé approximants are compared to a two-point Padé method which matches the c.d.f. around zero and infinity. The advantage of this technique is that only low order moments are required. This is most important when the moments are estimated, and higher order moments are significantly in error. The method is applied to some common problems in signal detection theory.

A Memoryless Grouped-Data Nonparametric Sequential Detection Procedure

M. M. Al-Ibrahim and P. K. Varshney *Department of Electrical and Computer Engineering, 111 Link Hall, Syracuse University, Syracuse, NY 13244-1240*

A nonparametric sequential testing procedure for the detection of constant signal in additive symmetric noise is proposed and analyzed. The nonparametric test is obtained by first quantizing all the observations into their signs, and then applying the Lee-Thomas modified sequential procedure to the quantized observations, or their sufficient statistic. The procedure is also extended to a distributed system consisting of two local detectors and a global decision maker. A simple truncation scheme is considered and is shown to maintain the nonparametric property while efficiently limiting the random test duration. Numerical results are presented to illustrated the performance gain and to demonstrate the simplicity of the test structure.

A Simple Approach to the Design of Decentralized Bayesian Detection Systems

W. Hashlamoun and P. K. Varshney *Dept. of Electrical & Computer Engineering, 111 Link Hall, Syracuse University, Syracuse, NY 13244-1240*

This paper considers the design of optimal decentralized Bayesian detection systems. A simple approach based on a version of Kolmogorov variational distance is presented. The design complexity of our approach is much less than the previous methods. A two sensor example is given for illustration.

Decision Agreement and Link Usage in Distributed Detection Systems with Feedback

Sam. Alhakeem, R. Srinivasan, and P. K. Varshney *Electrical and Computer Engineering Department, Syracuse University, 111 Link Hall, Syracuse, NY 13244-1240*

We consider a decentralized detection system with feedback. The issues of agreement probability and link usage in this feedback structure are investigated. Using some properties of this structure, we derive asymptotic results for the agreement probability. The second issue addresses the use of communication links between detector and the fusion center. We propose and analyze two protocols to

reduce the link usage. We derive equations for the average number of links needed under both hypothesis H_1 and H_0, and study the asymptotic performance.

A Converse Theorem for a Class of Multiterminal Detection Problems

Hossam M. H. Shalaby and Adrian Papamarcou *Electrical Engineering Department and Systems Research Center, University of Maryland, College Park, MD 20742*

We discuss the problem of testing a bivariate hypothesis $H : P_{XY}$ against a simple alternative $\bar{H} : \bar{P}_{XY}$ on the basis of i.i.d. pairs of discrete-valued observations (X,Y). We assume that the Y data are directly accessible to the decision maker or detector, while the X data are compressed at asymptotically zero rate. For $\bar{P}_{XY} > 0$, we show that the error exponent of the optimal test of level ϵ (i.e., the test that minimizes the type II error probability subject to the type I error probability not exceeding ϵ) is independent of ϵ for any $\epsilon \in (0,1)$. In doing so we generalize a previous result by Han that demonstrated the constancy of the error exponent for ϵ ranging in a subinterval of (0,1). Our result readily extends to hypothesis testing involving multivariate distributions on higher dimensional product spaces.

SESSION MA3

NEURAL NETWORKS I

Nested Neural Networks and Their Codes

Yoram Baram *Dept. of Elec. Engrg., Technion, Israel Institute of Technology, Haifa, 32000, Israel* (40 min.)

Neural networks can be viewed as encoders that employ sets of stored neural patterns as codes. Fully connected binary networks have been shown to provide low storage capacity and error correction capability. Networks composed of nested layers of small subnetworks are defined and shown to retrieve those permutations of the subcodes stored in the subnetworks that agree in their common bits. The resulting code sizes are considerably greater than those allowed by fully connected networks and the number of interneural connections is considerably smaller. Nested codes defined on subcodes consisting of two subwords are shown to be stable states of the network and to provide a relative error correction capability near 0.5 per subnetwork. Orthogonality of the subwords is shown to guarantee that their permutations are stable states of the nested network and that errors of relative size r per subnetwork will be corrected if the subcode stored in each subnetwork is of size smaller than $1/(2r)$. For randomly stored subcodes, the nesting property is shown to increase the probability of nondivergence and of error correction. In nondiscriminatory storage of such subcodes, stability and error correction capability can be guaranteed if and only if the subcodes stored in the subnetworks are of size 2 at most. Nested codes and their sizes are characterized for general subcodes and for specific orthogonal ones.

On the Number of Spurious Memories in the Hopfield Model

Jehoshua Bruck and Vwani P. Roychowdhury *IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120-6099 and Information Systems Laboratory, Stanford, CA 94305*

We show that the outer-product method for programming the Hopfield model can result in many spurious stable states--exponential in the number of vectors that we want to store--even in the case when the vectors are orthogonal.

Some Statistical Convergence Properties of Artificial Neural Networks

Andrew R. Barron *Departments of Statistics and Electrical & Computer Eng., University of Illinois, 725 S. Wright Street, Champaign, IL 61820*

Convergence properties of empirically estimated neural networks are examined. In this theory, an appropriate size feedforward network is automatically determined from the data. The networks we study include two and three layer networks with an increasing number of simple sigmoidal nodes, multiple layer polynomial networks, and networks with certain fixed structures but an increasing complexity in each unit. Each of these classes of networks is dense in the space of continuous functions on compact subsets of d-dimensional Euclidean space, with respect to the topology of uniform convergence. In this talk we show how, with the use of an appropriate complexity regularization criterion, the statistical risk of network estimators converges to zero as the sample size increases. Bounds on the rate of convergence are given in terms of an index of the approximation capability of the class of networks.

It's OK to be a *Bit* Neuron

Santosh S. Venkatesh *Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104*

We investigate computational and learning attributes in formal neurons when the neural weights are constrained to be binary (*bit* neurons). We demonstrate formally that there is little computational loss in eschewing real interconnections in favor of binary links: *with binary weights the achievable computational capacity is a much as one half that with real interconnections*. Analogous results hold for

learning weights from instances of a problem. Learning real weights for a McCulloch-Pitts neuron is equivalent to linear programming and can hence be done in polynomial time. Efficient local learning algorithms such as the Perceptron Training rule further guarantee convergence in finite time. The problem becomes considerably harder, however, when it is sought to learn binary weights; this is equivalent to integer programming which is known to be NP-complete. In the second part of this paper we present a new family of probabilistic binary learning algorithms. These algorithms have low computational demands and are essentially local in character. Rapid mean convergence times are demonstrated.

On Reliability and Capacity in Neural Computation

Santosh S. Venkatesh and Demetri Psaltis *Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104, and Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125*

We investigate the computing capabilities of formal McCulloch-Pitts neurons when errors are permitted in decisions. Specifically, we investigate the following:

Epsilon Capacity: In a fully interconnected recurrent network of formal neurons what is the maximum number of memories that can be stored given a prescribed level of recall error tolerance?

We show the following rigorous results given $0 \leq \varepsilon < 1/2$, a fractional error tolerance:

- Consider a random m-set of points in n-space and a set of associated binary decisions (or classifications) to be made on these points by a single formal neuron: the sequence $2n/(1-2\varepsilon)$ is a threshold function for the property that there exists a choice of synaptic weights for the neuron such that no more than (essentially) εm random decision errors is made; further, there is a function $1 \leq \kappa_\varepsilon < 505$ such that if m exceeds $2\kappa_\varepsilon n/1(1-2\varepsilon)$ then there is asymptotically no choice of synaptic weights for which a neuron makes fewer than εm decision errors on the m-set of inputs.
- Consider a random m-set of memories to be stored in a fully interconnected recurrent network of n neurons: the sequence $2n/(1-2\varepsilon)$ is a threshold function for the property that there is a choice of neural interconnections such that there are no more than (essentially) εn random bit errors in recall of any memory; if m is chosen larger than $2\kappa_\varepsilon n/(1-2\varepsilon)$ then for no choice of interconnections is any memory confined within a Hamming ball of radius εn.

For small ε the function κ_ε is close to 1 so that, informally, we can specify m-sets of points (memories) as large as $2n/(1-2\varepsilon)$ (but not larger) and obtain ε-reliable decisions (ε-error tolerance) for some suitable choice of synaptic weights.

Complexity of a Finite Precision Neural Network Classifier

K. Siu, A. Dembo, and T. Kailath *Information Systems Laboratory, Stanford University, Stanford, CA 94305*

A rigorous analysis of the finite precision computational aspects of a neural network as a pattern classifier via a probabilistic approach is presented. Even though there exist negative results on the capability of the perceptron, we show the following positive results: given n pattern vectors, each represented by cn bits where $c > 1$, that are uniformly distributed, with high probability the perceptron can perform all possible binary classifications of the patterns. Moreover, the resulting neural network requires a vanishingly small proportion $O(\log n/n)$ of the memory that would be required for complete storage of the patterns. Further, the perceptron algorithm takes $O(n^2)$ arithmetic operations with high probability, whereas other methods such as linear programming take $O(n^{3.5})$ in the worst case. We also indicate some mathematical connections with VLSI circuit testing and the theory of random matrices.

The Information Provided by a Linear Threshold Function with Binary Weights

Rodney M. Goodman, John W. Miller, and Padhraic Smyth *Dept. of Electrical Engineering (116-81), California Institute of Technology, Pasadena, CA 91125, and Communication Systems Research, Jet Propulsion Laboratories 238-420, 4800 Oak Grove Drive, Pasadena, CA 91109*

The *J* measure, originally used to measure the expected change in information provided by a production rule, is applied to measure the information provided by a linear threshold function (LTF) with binary (0,1) weights. This LTF with binary weights implements the logical function meaning "*X* of these *N* inputs are *ON*", called an *(X of N) rule*. For example a (1 of *N*) rule is an OR rule, and an (*N* of *N*) rule is an AND rule. In terms of the memory required to describe a basic logical unit, the (*X* of *N*) rule lies between the general LTF and the logical AND and OR functions. Logical AND and OR functions are used in production rule systems as a compact representation of domain knowledge which can be understood easily by humans. Efficient algorithms exist for searching through a database for informative conjunctive production rules. This paper extends this work to show how sets of informative (*X* of *N*) rules can be found, and shows why the (*X* of *N*) rules can be more powerful than conjunctive and disjunctive rules as unit of knowledge.

SESSION MA4

DATA COMPRESSION

Almost Sure Data Compression for Processes

Paul C. Shields *Dept. of Math., U. of Toledo, Toledo , OH 43606*

Universal coding for processes was first discussed in the landmark papers of Davisson, Lynch, and Fittinghof, and generalized by many subsequent authors to obtain asymptotic optimality in terms of expected value criteria, convergence in probability, or L^1 convergence. By extending to the rate-distortion setting a new entropy estimation technique of Ornstein and Weiss, ("How sampling reveals a process," Annals of Prob. (to appear)), we show how to construct a universal sequence of codes that is asymptotically optimal in the almost sure sense. The codes use the empirical distribution of nonoverlapping k-blocks in a sequence of length n; all that is is required is that $n \geq A^{2k}$, where A is the alphabet size. Some methods for computer implementation of these ideas will also be discussed. (This talk is based on joint work with D. S. Ornstein that will appear in the Annals of Probability.)

Finite Memory Modeling of Individual Sequences with Applications to State Estimation and Universal Data Compression

Marcelo J. Weinberger, Abraham Lempel, and Jacob Ziv *Technion - Israel Institute of Technology, Haifa 32000, Israel*

This paper deals with the estimation and the universal compression of discrete sources. The notion of stability of a word relative to an individual sequence over a finite alphabet is defined. In case of sequences emitted from a probabilistic finite memory source, the concept of stability is shown to be closely related to the estimation of the set of states of the source.

A Fast Construction Algorithm of Coding Tree for Variable-Length Data-Compression Coding with Fidelity Criterion

Hisashi Suzuki and Suguru Arimoto *Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan*

The problem of *variable-length data-compressing coding with* Δ*-distortion* aims to realize some code such that: as the message length approaches infinity, the expectation of the codeword length divided by message length approaches the Δ-distortion rate, and, the probability that the distortion between messages and their reproductions may not exceed Δ approaches 1. This paper proposes a top-down construction algorithm of balanced tree with a pointer indexed by an arbitrary-given distortion measure, and applies it to a method of variable-length data-compression coding. The main specifications of the obtained coding method are the following: the computation time for constructing a coding tree per codeword is $O(k)$, and that for each of encoding and decoding per codeword is also $O(k)$, where k denotes the message length; the expectation of the codeword length divided by message length approaches the Δ-distortion rate; the probability that the distortion between messages and their reproductions may not exceed Δ-distortion rate; the probability that the distortion between messages and their reproductions may not exceed Δ approaches 1.

Optimum Bit Allocation *via* the Generalized Breiman, Friedman, Olshen, and Stone Algorithm

Eve A. Riskin and Robert M. Gray *Information Systems Laboratory, Stanford University, Stanford, CA 94305*

An extension of Breiman, Friedman, Olshen, and Stone's algorithm for optimal pruning in tree-structured classification and regression is applied to bit allocation. It uses a simple tree model to

deallocate bits from classes of a coder, resulting in a set of variable rate codes of different average rates. When applied to classified vector quantization, it leads to a 3.4 dB gain in the signal-to-noise ratio of a magnetic resonance image coded at 1 bit per pixel over fixed rate classified vector quantization at the same rate.

Asymptotic Optimality of a Universal Variable-to-Fixed Length Binary Source Coding Algorithm

Tjalling J. Tjalkens and Frans M. J. Willems *Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

The modified Lawrence algorithm was introduced by the authors in 1988 as a universal binary variable-to-fixed length source coding algorithm. Here we investigate its asymptotic performance. For M (the segment set cardinality) large enough, we show that the rate $R(p)$ as function of the source parameter p satisfies

$$R(p) \leq h(p)\cdot(1 + \frac{\log(\log(M))}{2\cdot\log(M)})$$

for $0 < p < 1$, where $h(\cdot)$ is the binary entropy function.

In addition to this, we prove that no codes exist that have a better asymptotic performance, thereby establishing the asymptotic optimality of our modified Lawrence code.

The asymptotic bounds show that universal variable-to-fixed length codes can have a significantly lower redundancy than universal fixed-to-variable length codes with the same number of code words.

An Entropy Constrained Quantization Approach for a Source Characterized by a Random Parameter

Chein-I Chang and Lee D. Davisson *Department of Electrical Engineering, University of Maryland, Baltimore County Campus, Baltimore, MD 21228, and Electrical Engineering Department, University of Maryland, College Park, MD 20742*

The problem of quantizing a source characterized by a random parameter is considered. The method proposed in this paper is an entropy-constrained approach. The idea is that for a given source output partition we use a source matching algorithm to find the minimax entropy for the class of sources generated by the random parameter; then minimize the mean squared quantization error subject to the entropy constraint which is obtained by maximizing the minimax entropy over all possible partitions of the source output space. Since it is known that the uniform quantizer achieves the optimum performance at high-bit rates, asymptotic results for the proposed approach can be further obtained by replacing the optimum quantizer with the uniform quantizer.

The Redundancy Theorem and New Bounds on the Expected Length of the Huffman Code

Raymond W. Yeung *AT&T Bell Laboratories, Crawfords Corner Road, Holmdel, NJ 07733-1988*

We introduce the Redundancy Theorem as a tool to lower bound the expected length of prefix codes. It is shown that virtually all the previously known lower bounds of the expected length of the Huffman code can be obtained via applications of the Redundancy Theorem, and we demonstrate an application of the theorem which yields new lower bounds. We also obtain a new upper bound of the expected length of the Huffman code which depends on the entropy of the source and the two smallest probabilities of the distribution.

Alphabetic Codes Revisited

Raymond W. Yeung *AT&T Bell Laboratories, Crawfords Corner, Holmdel, NJ 07733-1988*

An alphabetic code for an ordered probability distribution $<p_k>$ is a prefix code in which p_k is assigned to the kth codeword of the coding tree in the left-to-right order. This class of codes is applied to binary test problems. We derive the characteristic inequality for alphabetic codes which is analogous to

the Kraft Inequality for prefix codes. With this inequality, we are able to unify and enhance many previous results on alphabetic codes. We discover that when $<p_k>$ is in ascending or descending order, the expected length of an optimal alphabetic code for $<p_k>$ is the same as that of a Huffman code for the unordered distribution $\{p_k\}$. We also prove two new lower bounds of the expected length of an optimal alphabetic code, and propose a simple method for constructing good alphabetic codes when optimality is not critical.

SESSION MA5

CHANNEL CAPACITY

On the Capacity of a Spectrally Constrained Poisson-Type Channel

Amos Lapidoth and S. Shamai (Shitz) *Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

The classical direct detection optical channel is modeled by an observed Poisson process of intensity $\lambda_t + \lambda_o$ wher λ_t is the information carrying waveform and λ_o stands for the dark current intensity. An upper bound on the capacity of this channel subjected to peak power $0 \le \lambda_t \le A$, average power $E(\lambda_t) = \sigma$ and covariance (spectral) constraints $V(\tau) = E(\lambda_t \lambda_{t+\tau}) - E^2(\lambda_t)$ is derived using the Kabanov-Davis approach that interrelates mutual information with nonlinear filtering in the Poisson regime. The resultant bound incorporates known results of optimal linear filtering in the Poisson regime and coincides with the known (peak and power constrained only) capacity as the spectral constraints are relaxed.

This new bound can be interpreted as the Poisson channel equivalent of Shannon's information integral $\frac{1}{2}\int_{-\infty}^{\infty} \log[1 + 2S_x(f)/N_o]df$ which upper bounds the capacity of the additive white Gaussian channel, where $S_x(f)$ and $N_o/2$ are respectively the input signal and noise power spectral density functions.

Lower bounds on the capacity under second moment and strict bandwidth constraints are found using for the modulating waveform λ_t, a special class of pulse amplitude modulated signals satisfying the required constraints.

Some Results on Zero-Error Capacity Under List Decoding

Erdal Arikan *Department of Electrical Engineering, Bilkent University, P.O. Box 8, Maltepe, Ankara 06572, Turkey*

We address a number of questions recently raised by Elias. Let $M_N(L)$ be the size of a maximal zero-error list-of-L code with blocklength N. Let $C_0(L)$ be the zero-error list-of-L capacity defined by

$$C_0(L) = \liminf_{N\to\infty} \frac{1}{N} \log M_N(L)$$

Elias shows that

$$\lim_{L\to\infty} C_0(L) = C_{0F}$$

where C_{0F} is the zero-error capacity of the same channel with feedback. We prove the following results:

1. Computing C_{0F} is an *NP*-complete problem.
2. For all $L \ge 1$ and $N \ge 1$,

   $$M_{N+1}(L) \le \lfloor 2^{C_{0F}} M_N(L) \rfloor$$

   >From this we obtain, e.g., that, for the 3/2-channel discussed by Elias, $M_4(2) \le 9$, answering one of his questions.
3. For all $L \ge 1$, $N \ge 1$, and $1 \le K \le L$,

   $$M_N(L) \ge M_{N+n}(K)$$

   where n is the smallest integer such that $M_n(K) \ge L$.

The question of how tight these bounds are is also discussed.

The Capacity-Cost Function of a Noiseless Channel with Several Cost Constraints

Robert J. McEliece and Lada Popović *Department of Electrical Engineering, California Institute of Technology, MC 116-81, Pasadena, CA 91125*

A noiseless channel is usually represented by a directed graph whose vertices are the channel symbols, the allowable input sequences being the walks in the graph. Some properties of the channel can be modeled by assigning to each edge a vector of $\mathbf{c} \in \mathbf{R}^m$ and requiring that only those walks are allowed whose cost-per-edge does not exceed, componentwise, some given vector $\mathbf{w}$.

We define the *Capacity-cost function*

$$C(\mathbf{w}) = \limsup_{n \to \infty} \frac{\log M(n, \mathbf{w})}{n}$$

(where $M(n, \mathbf{w})$ is the number of walks of length n and per-edge cost $\leq \mathbf{w}$) and show that it is equal to the maximal entropy of a Markov chain which is supported by the graph and whose expected cost per transition does not exceed $\mathbf{w}$. We also derive an expression for $C(\mathbf{w})$, thus extending the result of Justesen and Hoholdt to multidimensional costs.

Using the above results, we obtain a formula for the exponential decay rate

$$\limsup_{n \to \infty} \frac{\log P\{K(x) \leq np\}}{n}$$

of the probability that a sequence has less than a specified per-transition cost on a given Markov chain with real costs assigned to the transitions.

A New Upper Bound on ϵ-Capacity

Michael L. Honig and Prakash Narayan *Bellcore, 445 South Street, Morristown, NJ 07960 and Department of Electrical Engineering and Systems Research Center, University of Maryland, College Park, MD 20742*

Suppose that two outputs of a linear, time-invariant channel are distinguishable at the receiver if and only if they are separated in L_2 norm by ϵ. The inputs to the channel are assume to be power limited, and are nonzero only on the finite time interval $[0, T]$. Let $N_{\max}(T, \epsilon)$ be the maximum number of distinguishable outputs for given T, $\epsilon > 0$. The ϵ-capacity of the channel is defined as $C_\epsilon = \lim_{T \to \infty} \log_2[N_{\max}(T, \epsilon)]/T$ bits/second. It has been shown by Forys and Varaiya that C_ϵ is upper bounded by the Shannon capacity, C_S, of a channel consisting of the original channel followed by additive Gaussian noise with variance $\epsilon^2/4$, but otherwise having arbitrary spectral density. We show that the noise spectral density that minimizes C_S is proportional to the power spectral density of the input. It is also shown that the resulting upper bound on C_ϵ is less than or equal to the Shannon capacity of a channel consisting of the original channel plus an additive noise source (not necessarily Gaussian) with variance $\epsilon^2/4$, but otherwise having arbitrary statistics. Numerical results are presented for models of subscriber loop channels that show the new upper bound is significantly better than the upper bound previously derived by Root.

A Lower Bound on the Capacity of Primitive Binary BCH Codes Used in Gaussian Channel with Discrete Time

Dejan E. Lazić and Vojin Šenk *Universität Karlsruhe, Institut für Algorithmen und Kognitive Systeme, D-7500, Karlsruhe 1, Fed. Rep. of Germany, and University of Novi Sad, Computer Science, Control and Measurements Institute, Novi Sad, Yugoslavia*

Since the BCH codes are not asymptotically good, it was long believed that long BCH codes do not behave well in a communication channel. This conjecture was based on the assumption that the ratio of the minimum Hamming distance of the code to the code length exhibits the dominant influence on the error probability for any signal-to-noise ratio and any dimension (code length). This conjecture is

disproved in this paper, showing that in the additive white Gaussian noise channel, for large enough dimension and low signal-to-noise ratio SNR the Hamming distance and the Euclidean distance that is monotonically related to it are of no influence on the error probability, provided that the multiplicity of code words on the distance is below a certain bound. The capacity of long primitive binary BCH codes used in Gaussian channel with discrete time is, using this result, lower-bounded by

$$R_{BCH} = 1 - ld\,(1 + 2^{-SNR/2\ln 2})$$

i.e., for a given SNR very long primitive binary BCH codes of rate R_{BCH} can be safely used in the Gaussian channel, provided that true minimum-distance decoding is used instead of the usual hard-decision bounded-distance decoding method.

Zero-Error Capacities and Very Different Sequences

G. Cohen, J. Körner,and G. Simonyi *ENST, 46, rue Barrault 75634 Paris Cedex 13, France, and Mathematics Institute of HAS, 1364 Budapest, POB 127, Hungary*

Perfect hash functions, superimposed codes as well as some other fashionable questions in computer science and random-access communication are special cases of early-day information theoretic models in the zero-error case.

A new class of problems in asymptotic combinatorics can be formulated as the determination of the zero-error capacity of a class of discrete memoryless channels. (This model is also known as the compound channel). We solve an interesting class of these problems using our recent results in polyhedral combinatorics.

Capacity of the Gaussian Arbitrarily Varying Channel

Imre Csiszár and Prakash Narayan *Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, POB 127, Hungary, and Electrical Engineering Department and the Systems Research Center, University of Maryland, College Park, MD 20742*

The Gaussian arbitrarily varying channel with input constraint Γ and state constraint Λ admits input sequences $\mathbf{x} = (x_1, \ldots, x_n)$ of real numbers with $1/n \sum x_i^2 \le \Gamma$ and state sequences $\mathbf{s} = (s_1, \ldots, s_n)$ of real numbers with $1/n \sum s_i^2 \le \Lambda$, the output sequence being $\mathbf{x} + \mathbf{s} + \mathbf{V}$ where $\mathbf{V} = (V_1, \ldots, V_n)$ is a sequence of independent and identically distributed Gaussian random variables with mean 0 and variance σ^2. We prove that the capacity of this arbitrarily varying channel for deterministic codes and the average probability of error criterion equals $\frac{1}{2} \log (1 + \frac{\Gamma}{\Lambda + \sigma^2})$ if $\Lambda < \Gamma$ and is 0 otherwise.

SESSION MA6

CODING THEORY I

Exponential Error Bounds for Randomly Modulated Codes on Gaussian Channels with Arbitrarily Varying Interference

Brian Hughes and Tony G. Thomas *Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218* (40 min.)

The Gaussian arbitrarily varying channel (GAVC) models a channel corrupted by thermal noise and by an unknown interfering signal of bounded power. In this paper, we present upper and lower bounds to the best error probability achievable on this channel with random coding. The asymptotic exponents of these bounds agree in a range of rates near capacity and at rate zero. The exponents are *universally larger* than the corresponding exponents for the discrete-time Gaussian channel with the same capacity.

We further show that the optimal exponents can be achieved by a restricted kind of random code comprised of a deterministic encoder/decoder in cascade with an independent linear random modulator/demodulator. Moreover, the decoder can be taken to be the minimum Euclidean distance rule at all rates less than capacity. Connections to spread-spectrum modulation are discussed.

Spectral Lines of Codes Given As Functions of Finite Markov Chains

Hiroshi Kamabe *Department of Electronics, Mie University, Tsu 514, Japan*

We analyze the spectral line of a memoryless function of a finite Markov chain. We expect that the results of this paper will have applications in fiber optics. A memoryless function of a finite Markov chain is given by a triple (G,P,γ), where G is a directed graph, P is a transition probability matrix compatible with G and γ is a complex valued function of the vertex-set of G. We prove that for G, γ and a positive real number c the following two conditions are equivalent: (1) for any P the spectral line of (G,P,γ) at a rational submultiple f of the symbol frequency is equal to (not less than, respectively) c; (2) for any loop of G, if its length is L, then the absolute value of the running digital sum of the sequence of values of γ of the vertices along the loop at f, is equal to (not less than, respectively $L\sqrt{c}$. We also characterize these conditions in terms of coboundary functions introduced by B. Marcus and P. Siegel.

On the Construction of Statistically Synchronizable Codes

R. M. Capocelli, A. De Santis, L. Gargano, and U. Vaccaro *Dipartimento di Matematica, Universita' di Roma, 00185 Roma, Italy; IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York, 10598; and Dipartimento di Informatica ed Applicazioni, Universita' di Salerno, 84100 Salerno, Italy*

We consider the construction of almost-optimal statistically synchronizable codes for arbitrary alphabets and finite sources. We show that it is always possible to obtain a code possessing a synchronizing sequence and presenting an increase on the optimal average code word length not greater than the lowest probability associated to a codeword. Moreover, we give an efficient method to construct codes having a synchronizing word. These codes are almost-optimal, in the sense of a small average length, and present high synchronization capability. All previous proposed solutions were restricted to particular sources and considered only a binary alphabet.

'1'-Ended Binary Prefix Codes

Renato M. Capocelli and Alfredo De Santis *Dipartimento di Matematica, Universita' di Roma, 00185 Roma, Italy and IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York, 10598*

Binary prefix codes with the constraint that each codeword must end with a '1' have been recently introduced by Berger and Yeung. We analyze the performances of such codes by investigating their average codeword length. In particular, we show that a very simple strategy permits the construction of a '1'-ended binary prefix code whose average codeword length is within $H + 1$ for any discrete source with entropy H. We also prove a tight lower bound on the optimal average codeword length in terms of H and the minimum probability of the source. Finally, we discuss the problem of finding an optimum code.

Mathematical Models for Block Code Error Control Systems on Renewal Inner Channels

D. R. Oosthuizen, H. C. Ferreira, and F. Swarts *SA Transport Services, Telecommunication and Laboratory for Cybernetics, Rand Afrikaans University, PO Box 524, Johannesburg 2000, South Africa*

We investigate the average event rate and statistical distribution of the block error detection, correction and misdetection events of block codes on discrete renewal inner channels, when the latter are represented with partitioned Markov chains as proposed by Fritchman. We show that the underlying statistical structure of the events under consideration can be represented with outer channel models, which are also partitioned Markov chains, similar to the inner channel models. Analytical procedures to determine the parameters of such an outer channel model, as a function of the inner channel model parameters, are presented. The results of extensive computer simulations to verify the analytical procedures, are also presented. The application of the outer channel models in the design of error control systems, such as systems with repetition of codewords and rejection of codewords with detected errors, as well as ARQ systems, is investigated.

Run Length Coding with Spectral Lines

Kenneth J. Kerpez *Bell Communications Research, Morristorn, NJ 07962-1901*

Codes are considered that map information into a binary run length limited sequence with spectral lines. The codes provide pilot tones for tracking in magnetic or optical recording. The coded sequences are generated by variable length state transition diagrams that are periodic with period p. A sufficient condition that the code has spectral lines is that the square of the state transition diagram has period $2p$. The poles in the spectrum on the unit circle occur at the p^{th} roots of unity; the lines occur at the corresponding frequencies. A deterministic component in the coded sequence has a periodic autocorrelation, which transforms to a line spectrum. A method of computing the power in the line spectrum from stationary probabilities of the state diagram is shown. Relations are given that prove that as the power in the line spectrum or the number of lines increases, the information capacity of the code decreases. Two families of graphs are shown that produce codes with spectral lines.

Error Free Coding on Chinese Characters

Rong-Hauh Ju, I-Chang Jou, Mo-King Tsay, and Kuang-Yao Chang *Telecommunication Laboratories, Ministry of Communications, P.O. 71, Chung-Li, Taiwan, R.O.C., and Institute of Information & Electronics, National Central University Chung-Li, Taiwan, R.O.C.*

Dot pattern representation is the most popular method to represent Chinese characters. However, the storaging memory of these Chinese character patterns is too large to be processed well. In this paper, the characteristics of Chinese patterns are introduced and analyzed first, then some data-compression methods have been tried, aiming at the smallest amount of data storage. The compression algorithms (preprocessing, statistics of source symbols, encoding), which are described by three passes, will be discussed by their compression ratio, entropy, and complexity. We find out that the compression method (prediction, subblock, arithmetric coding) will achieve a best compression ratio of about 25% and save

about 75% of the storaging memory. Besides, the entropy of Chinese characters, which is preprocessed by prediction, is almost independent of resolution of the subblock. This property gives us a hint to understand a possible bound on compression on Chinese characters with various resolutions.

Random Modulation/Coding Problems for a General Channel

Shi yi Shen *Department of Mathematics, Nankai University, Tianjin, P.R. China*

A modulation/coding system (MCS) is said to be a random MCS (RMCS) if the input constellation of channel is a random constellation. A channel is said to be a general channel if its noise is a general noise or the input and output signals are general random variables (Gaussian or non-Gaussian, discrete or non-discrete). In this paper, we first obtain some asymptotic formulas for the channel capacity of the RMCS. Then we give a general discussion of the coding problems for the RMCS under block codes, linear codes, and trellis codes. Some coding theorems are proved, and an error-bound theorem is obtained for the Gaussian channel. These results show some basic characteristic for the MCS and general channels

SESSION MA7

TRELLIS CODING I

Coherent and Differentially Coherent Trellis Coded Modulation on Channels with Correlated Time-Selective Fading

Christian Schlegel *Communications Group, ASEA BROWN BOVERI Corporate Research, 5405 Baden, Switzerland*

Trellis-coded modulation (TCM) on correlated fading channels, described by the wide sense stationary uncorrelated scatterer (WSSUS) model, is studied. For this channel model the two code error probability can be expressed as the probability that certain Gaussian quadratic forms exceed zero. It is shown how the event error probability P_e, upper bounded by the averaged union bound, can be calculated with an efficient algorithm using quasi-regularity properties of the codes used. It is further shown that in the limiting case of no interleaving, the Euclidean distance spectrum can be used to evaluate P_e, while for full interleaving, the effective length/minimum product spectrum determines performance.

It is further shown how differential TCM on correlated fading channels, in conjunction with the suboptimum Euclidean distance metric, can be treated analytically. Results show that differential TCM suffers a degradation between 2.5 dB and 5 dB with respect to coherent TCM.

In a refinement of the analysis, non-ideal estimation of the channel impulse response in the coherent case and co-channel as well as adjacent channel interference are included.

Analysis of the Error Performance of Trellis-Coded Modulations in Rayleigh Fading Channels

Jim Cavers and Paul Ho *School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6*

Trellis-Coded Modulation (TCM), when combined with interleaving, is known to give good error performance in fading channels. Previously, though, the only analytical guide has been an upper bound on the pairwise error event probability, which could be very loose over the range of signal to noise ratio of interest. In contrast, this paper presents an **exact expression** for the pairwise error event probability of TCM transmitted over Rayleigh fading channels. It includes PSK and multilevel QAM codes, as well as coherent and partially coherent (e.g., differential, pilot tone, etc.) detection. We have found that a good estimate of the bit error probability can be obtained by considering only a small number of short error events. This implies digital computer simulation can be avoided during the link design process. We study several coded modulation schemes this way. Among the results are the fact that compared to uncoded modulation, TCM provides a significant improvement in the error floor when detected differentially, and an asymmetry in the pairwise error event probability for 16QAM.

Trellis Coded MPSK Modulation with an Unexpanded Signal Set

Shalini S. Periyalwar and S. Fleisher *Department of Electrical Engineering, Technical University of Nova Scotia, Halifax, NS B3J 2X4*

In this paper, a scheme for transmitting two channel symbols per trellis branch is proposed, wherein the output bits from a rate $m/m+1$ encoder are used to select *pairs* of symbols from a 2^m-ary (unexpanded) MPSK symbol set. The proposed set partitioning technique *simultaneously optimizes* the symbol assignments to the trellis branches for performance on the AWGN channel (maximize $d^2(free)$) and the fading channel (maximize diversity and product of branch distances along error event path of shortest length). The throughput rate is $m/2$ bps/Hz. For a given throughput rate, the complexity of the scheme (measured by symbol multiplicity, number of parallel paths, and number of states) lies between that of the TCM and MTCM schemes, and the performance gains achieved on the AWGN and fading channels are equal to or higher than those of MTCM schemes. We demonstrate results for MPSK

schemes with throughput 1.0 bps/Hz, 1.5 bps/Hz and 2.0 bps/Hz. For example, the rate 3/4 scheme (throughput 1.5 bps/Hz) achieves the same $d^2(free)$ (= 8.0) for 4 states, as the rate 6/12 MTCM scheme (with set partitioning optimized for the AWGN channel), with lower complexity (multiplicity of symbols: (2 *vs.* 4), number of parallel paths: (2 *vs.* 8)).

This technique may be useful for transmission on the Rician fading channel with large values of the fading parameter, where the requirement for the simultaneous optimization of criteria for AWGN and fading channels cannot be achieved by set partitioning of MTCM for fading channels.

Fractional-Bit Transmission with Single-Symbol TCM

E. Eleftheriou and P. R. Chevillat *IBM Research Division, Zürich Research Laboratory, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland*

Exploiting the bandwidth of a transmission channel to the largest possible extent is often only possible if a non-integer number of bits per symbol is sent. In this paper we consider the transmission of TCM-coded signals with $m + p/q$ bits/symbol where $q = 2^k$. We employ single-symbol TCM codes together with simple block-coding schemes and signal constellations whose size is not a power of two. This combination allows to match the symbol rate to the available channel bandwidth while retaining the advantages of single-symbol TCM, e.g., a small decoding delay, and the availability of zero-delay tentative decisions for decision-directed receiver adaptation. In addition, the block code has the property that symbols with large amplitudes are sent less frequently improving performance in the presence of nonlinear distortion. The block-coding operations become particularly simple for $p = 1$. As an example, the transmission of 5.25 bits/symbol is discussed which offers an alternative to the CCITT V.33 scheme for 14.4 kbits/s modems. Finally, the concept of reduced-state combined equalization and trellis decoding which mitigates using a larger bandwidth is generalized to the case of fractional-bit transmission.

Advanced Synchronization Procedures for Trellis Coded MFSK Modems

B. Honary, F. Zolghadr, and M. Darnell *Department of Engineering, University of Warwick, Coventry, CV4 7AL, UK, and Department of Electronic Eng., University of Hull, Hull, HU6 7RX, UK*

Two of the major problems encountered with communication channels are those of channel errors and bit (or symbol) synchronization. Errors are normally countered by the use of appropriate error control coding. As a separate consideration the problem of synchronization is typically solved by an initial synchronization process, followed by either precise subsequent timing at the receiver, or use of segmenting information inserted within the transmission format. Although these procedures are not optimum, until recently they have represented the most efficient practical approach possible.

With the advent of powerful and cheap digital signal processing systems, however, it is now feasible to unify the two processes, thus potentially producing a more efficient communication system. In this presentation an example of such communication system is described. The system employs a new symbol synchronization method applicable to the reception of trellis coded MFSK signals. The procedure is termed code-assisted but synchronization (CABS), in which digital processing techniques are used to combine the functions of demodulation and error control decoding.

The performance of the CABS modem has been studied using simulation and practical techniques, under additive white Gaussian noise channel condition. It is shown that the degradation in performance (compared with a perfectly synchronized modem) is less than 1dB.

A Trellis Partitioning Technique for Reduced-State Sequence Detection

Torbjörn Larsson *Telecommunication Theory Group, Dept. of Information Theory, Chalmers University of Technology, S-412 96 Göteborg, Sweden*

Recently, several related algorithms for reduced-state sequence detection have been proposed by various authors. A joint description of these algorithms can be given by observing that they are all based

on a partitioning of the trellis, dividing the set of states into a number (C) of state-classes. The detector searches the trellis by saving the K best candidate paths from each state-class. Each choice of state-classes results in a new search algorithm with an associated minimum distance d_{min}. This minimum distance will, especially for $K = 1$, limit the detection performance achieved by the algorithm. Thus, it is of interest to determine the particular set of state-classes that maximizes d_{min}.

In this paper a systematic search procedure is employed to find the best (maximal d_{min}) partitioning of the trellis. As an application, we consider uncoded quadrature amplitude modulation with finite intersymbol interference (ISI). For every ISI channel there is a number C^* such that if $C \geq C^*$, then a partitioning with C state-classes can be found with d_{min} equal to the free distance of the channel. For minimum phase channels, it is found that C^* is usually only a fraction of the number of states in the trellis.

Soft Decision Demodulation and Multi-Dimensional Trellis Coded Phase Modulation

Joseph M. Nowack and Mark A. Herro *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556*

For a multi-dimensional trellis coded phase modulation system, quantization at the receiver will cause some performance degradation relative to infinitely fine quantization. In this paper, we study a variety of soft decision schemes that are designed to reduce the effect of quantization at a reasonable cost. The schemes are compared using cutoff rate calculations, simulations of multi-dimensional trellis coded phase modulation systems, and implementation considerations. In addition, several soft decision schemes having central erasure regions are compared based on cutoff rate calculations to study the influence of these erasure regions.

Distance Weight Distribution of Trellis Codes Found by DFT

Torleiv Maseng *Elab-Runit, N-7034 Trondheim, Norway*

A closed form analytic algebraic formula is given for the distance weight distribution of a trellis code. This has been done by the application of the transition matrix approach. The method uses characteristic functions in combination with a discrete Fourier transform rather than generating functions in combination with differentiation. This method could be faster to program and could also benefit from commonly available special hardware or software in order to do the Fourier transforms.

TECHNICAL SESSIONS

Monday, 2 p.m. - 5 p.m.

SESSION MP1

COMMUNICATION SYSTEMS

A Parallel Systems Approach to Universal Receivers

Upamanyu Madhow and Michael B. Pursley *Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Ave., Urbana, IL 61801*

We introduce a universal approach to dealing with uncertainty in channel characteristics. A parallel implementation is proposed for a universal receiver that can cope with such uncertainty. The parallel implementation consists of a finite number of receivers with the property that, for any channel in the class of interest, the performance of at least one of the receivers will be within a specified degradation of the optimal performance. The identification of the good receivers is accomplished by means of the intrinsic side information generated by an appropriate coding scheme. In this paper, we give sufficient conditions on channel classes for which a universal design as described above is possible. We also outline procedures for carrying out such a design, and give a general example for M-ary signaling to illustrate the applicability of our theory. Tools for the analysis of the coded performance of our parallel implementation are developed, and it is shown that Reed-Solomon codes with bounded distance decoding satisfy the requirements of a good coding scheme for our application. (Research supported by the Army Research Office (DAAL-03-87-K-0097).)

Timing Recovery in the ISDN U-Interface Transceiver

Erdal Panayirci *Faculty of Electrical and Electronics Engineering, Istanbul Technical University, Ayazağa Kampüsü, Istanbul, Turkey, and TELETAS R&D Department, Umraniye, Istanbul, Turkey*

In this paper, a general analysis is presented for the jitter performance of a common Symbol timing Recovery (STR) system employed in a digital subscriber Loop (DSL) transceiver employing adaptive echo cancellation for high-speed digital communications typical of evolving Integrated Service Digital Networks (ISDN's). The STR circuit consists of a prefilter and a squarer followed by a narrow-band postfilter tuned to the signalling rate. The output of the timing circuit is a nearby sinusoidal wave whose zero crossings indicate the appropriate sampling instants for extraction of the data. Exact analytical expressions for the mean and variance of the timing wave and for the rms phase jitter are derived as a function of the bandwidth of the postfilter for a given set of input parameters representing a particular digital subscriber loop and its noise environment, including such effects as residual-echo, crosstalk and impulse noise. Numerical results, obtained for an experimental study of a 144 kbit/s DSL timing recovery system show that the presence of these disturbing signals can substantially degrade the synchronizer performance. The effect of the excess bandwidth factor of the prefilter on this degradations is also investigated.

Some New Results and Interpretations Concerning Binary Orthogonal Signaling Over the Gaussian Channel with Unknown Phase/Fading

Pooi Yuen Kam *Department of Electrical Engineering, National University of Singapore, Kent Ridge 0511, Republic of Singapore*

We consider the well-known problem of binary orthogonal signaling over the Gaussian noise channel with unknown phase/fading. By viewing the problem in a rotated coordinate system, the

orthogonal signal structure can be considered as the combination of an antipodal signal set and an unmodulated component (pilot tone) for channel measurement. This allows us to show that as far as data detection is concerned the optimum matched-filter-envelope-detector is identical to a novel detector-estimator receiver in which the detector performs partially coherent detection using an absolute coherent reference generated by the estimator from the channel measurement provided by the pilot-tone component of the orthogonal signal structure. This detector-estimator interpretation of the optimum receiver shows that it is wrong to refer to the latter as a noncoherent receiver. It also leads to the development of new approaches for analyzing the error probability performance of the receiver. In particular, for the Rician channel, an exponential Chernoff upper bound is obtained, and an expression is also obtained for the case of slight fading.

Multidimensional Signaling with Parallel Architectures

E. Biglieri and F. Pollara *Electrical Engineering Department, UCLA, Los Angeles, CA 90024, and Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA*

Imai and Hirakawa and Ginzburg have shown how algebraic codes of increasing Hamming distance can be combined with nested signal constellations of decreasing Euclidean distance to generate multidimensional signals with large distances. Cusack and Sayegh have shown specific constructions of multidimensional constellations.

Staged demodulation of these multidimensional constellation was also suggested. With this procedure the bits of the signal label protected by the most powerful code are decoded first by using maximum-likelihood (soft) decoding. Then the bits protected by the second most powerful code are decoded, and so on. This procedure is suboptimum, i.e., the error probability will be increased, but reduced decoding complexity will result. In particular, staged decoding is amenable to an architecture with pipelined parallelism.

In this paper we consider the design of constellations that allow a high degree of parallelism in the staged decoder structure, and in addition allow soft decoding of the component algebraic codes by using a systolic algorithm suitable for VLSI implementation.

The resulting receiver structure turns out to be highly modular and flexible, so that different codes and different constellations can be accommodated. Finally, combination of these multidimensional constellations with Trellis-Coded Modulation (TCM) is considered, and it is shown that this highly parallel demodulator structure can work also if a TCM scheme is used to further increase the transmission performance.

Analysis of SCCL As a PN Code Tracking Loop

Kwang-Cheng Chen and Lee D. Davisson *Electrical Engineering Department, University of Maryland, College Park, Maryland 20742*

Synchronization is an essential issue for PN (pseudo noise) coded spread spectrum communication systems. Recent research in this field has concentrated on PN code acquisition. DLL (delay-locked loop) and TDL (tau-dither loop) are still two primary types for PN code tracking. Both of the loops are based on the design of the early-late gate bit synchronizer. Recently, Chen and Davisson proposed a new bit tracking loop, SCCL. Their design is a digital tracking loop that is based on the Sample-Correlate-Choose-Largest procedure. In this paper, we introduce the application of SCCL as a PN code tracking loop and analyze its performance both in steady-state and transient conditions. SCCL applying a biphase-level-level signal set acts as a coherent baseband PN code tracking loop which more closely implements the *MAP* optimal structure. Three adjacent estimates are formed by correlating the samples of the baseband waveform for each bit. We choose the corresponding timing (phase) of the estimate with the largest magnitude as the current correct timing (phase) and update it for each bit. Only one summation circuit is necessary due to the digital realization of the SCCL. The correlation properties of the samples from maximum length codes using the biphase-level-level signal set are investigated. A finite-state Markov chain model is used for theoretical performance analysis. Its transition probabilities can be

represented via triple integrals and calculated numerically. The numerical results of performance analysis are presented in this paper.

Adaptive Rate Sampling for Secure Communication Systems

M. Darnell and B. Honary *Department of Engineering, University of Warwick, Coventry, CV4 7AL, UK, and Department of Electronic Eng., University of Hull, Hull, HU6 7RX, UK*

The paper first considers a procedure for spectral manipulation, based upon a time domain sampling algorithm, applied to signals having spectra comprising multiple (> 2) separated band-pass elements. The algorithm provides an analytical and practical tool to enable spectral manipulation to be carried out in systematic manner, allowing non-uniform adaptive sampling rates to be applied to a multiple band-pass analog signal if the band structure is first characterized. Time variation in the analog signal structure can also be tracked. Given that the means are available to identify the form of a spectrum comprising multiple band-pass elements in terms of the transmission frequencies between each of the elements and the adjoining spectral 'gaps', the procedure can be used to drive the minimum sampling rate necessary to characterize that spectrum completely, and hence allow its complete recovery from the samples taken at that calculated rate. This sampling algorithm is then applied to the scrambling and recombination of a multiple band-pass spectrum. The procedure enables the separated band-pass elements of a frequency-scrambled signal, some of which may also be frequency-inverted, to be recombined in their correct order and possibly translated in frequency, if required. This would enable the scrambled band-pass elements of, say, a speech signal to be recombined into an intelligible speech signal. It thus provides the basis of a speech or data privacy/security communication system, in which the secure "key" information comprises (a) the spectral frequency ranges of the scrambled elements; and (b) the unique value of sampling rate which will allow correct recombination.

Study of Self-Noise Spectra in Fourth-Power Law Clock Recovery

Thomas T. Fang *Lockheed Missiles & Space Company, Inc., Orgn. 91-50, Bldg. 251, Palo Alto, CA 94304*

A prior work has established that average jitter power in a clock recovery circuit is a function of the self-noise powers in phase with the desired recovered clock and in phase quadrature as well as of the cross-spectrum between these two components. In this paper, we found a method to compute these spectral components for the fourth-power clock recovery technique for the PAM case. Comparison is made with the squaring type recovery technique using cosine roll-off Nyquist pulses on BPSK mode, at several excess bandwidth factors (α). A prefilter is described which completely eliminates time jitter in the fourth-power clock recovery circuit.

Nonlinear Self-Training Adaptive Equalization for Multilevel Partial-Response Class-IV Systems

Giovanni Cherubini *IBM Research Division, Zürich Research Laboratory, Säumerstr. 4, CH-8803 Rüschlikon, Switzerland*

Self-training adaptive equalization for multilevel partial-response class-IV systems is addressed. An adaptive equalizer realized with distributed-arithmetic architecture is considered, where the process of multiplying the tap signals with tap gains and summing the resulting product is replaced by a procedure involving only table look-up values and shift-and-add operations. Self-training adaptation schemes devised for linear adaptive equalizers do not converge if applied to a distributed-arithmetic equalizer, because of the inherent non-linearity of the system during the adaptation process. In this paper a new algorithm allowing the convergence of the look-up values to the optimum setting is proposed and the analysis of the dynamics of the look-up values is conducted. Under the usual assumption of independent

signal vectors in the equalizer delay line at different updating times, a sufficient condition allowing the temporal evolution of the look-up values to be modeled as an ergodic Markov process is given. Numerical results are presented with reference to a multi-level PRIV system for high-rate data transmission over twisted-pair cables.

SESSION MP2

BROADCAST CHANNELS

Selective Repeat ARQ Schemes for Broadcast Links

S. Ram Chandran and Shu Lin *Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822*

In this paper we have proposed several selective repeat ARQ schemes for point to multi-point communications, where the receivers have finite buffer size. The schemes proposed here can be classified into two categories: the hybrid schemes and the ARQ schemes. The ARQ schemes presented here reduce to special cases of the hybrid schemes when the forward error correction part is removed from them. We obtain lower bounds on the throughput efficiency of the schemes. Computation of the bounds reveal that the schemes perform quite satisfactorily for channels with large round-trip delay and high data rate like the satellite channel. All the schemes are adaptive and transmit multiple copies of the message at different levels of transmission. The schemes are simple to implement and outperform the ones proposed by Towsley and Mithal and by Mohan, Qian, and Rao. We have also applied the dynamic programming approach of Wang and Silvester to our schemes to choose the optimum number of copies. The resulting schemes give some improvement in throughput.

Identification for a Deterministic Broadcast Channel

Bart Verboven and Edward C. van der Meulen *Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3030 Leuven, Belgium*

We derive a direct identification result for the deterministic broadcast channel (BC). We also establish a 'soft converse' in the sense of Ahlswede and Dueck (1989), thereby solving the identification problem for this BC.

Surprisingly, for the deterministic BC the region of achievable ('second order') identification rates is much larger than the region of achievable transmission rates, in contrast to the one-way channel result obtained by Ahlswede and Dueck. More precisely, for the deterministic BC, all the second order rate pairs (R_1, R_2) satisfying

$$0 \le R_1 \le H(Y_1),$$

$$0 \le R_2 \le H(Y_2),$$

where (Y_1, Y_2) is the channel output corresponding to some input random variable X, are achievable.

As one can see, in this characterization of the identification capacity region no constraint is needed on the sum $R_1 + R_2$ of both rates, whereas there appears such an extra constraint in the description of the transmission capacity region obtained by Pinsker (1978) for this BC.

Communicating Via a Processing Broadcast Satellite

F. M. J. Willems, J. K. Wolf, and A. D. Wyner *Dept. of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands; Center for Magnetic Recording Research, University of California, San Diego, La Jolla, CA 92093; and AT&T Bell Laboratories, Murray Hill, NJ 07974*

Three dependent users are physically separated but communicate with each other via a satellite. Each user generates data which it stores locally. In addition, each user sends a message to the satellite. The satellite processes the messages received from the users and broadcasts one common message to all three users. Each user must be capable of reconstructing the data of the other two users based upon the broadcast message and its own stored data. Our problem is to determine the minimum amount of information which must be transmitted to and from the satellite.

Let the data of the three users at time n, $(1 \le n < \infty)$ be (X_n, Y_n, Z_n), which are statistically independent drawings of the dependent vector (X, Y, Z). Let R_x, R_y, and R_z be the rates of transmission up to the satellite by the three users, and let T_0 be the rate of transmission down from the satellite. We show that for data blocks of size N, where N is large, the set of rates (R_x, R_y, R_z, R_0) which satisfy

$$R_x > H(X \mid Y, Z), \quad R_y > H(Y \mid X, Z), \quad R_z > H(Z, X, Y),$$
$$R_x + R_y > H(X, Y \mid Z), \quad R_y + R_z > H(Y, Z \mid X), \quad R_x + R_z > H(X, Z \mid Y),$$
$$R_0 > \max[H(X, Y \mid Z), H(X, Z \mid Y), H(Y, Z \mid X)].$$

is necessary and sufficient for the purpose outlined above.

Some Matching Results in Multi-User Communication

Sergei I. Gelfand and Edward C. van der Meulen *Institute for Problems of Information Transmission, USSR Academy of Sciences, 19, Ermolova Street, 101447, GSP-4 Moscow, USSR, and Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, B-3030 Leuven, Belgium*

We derive necessary and sufficient conditions for the reliable transmission of a two-component correlated source over a multi-user channel in two situations. First, matching conditions are derived for sending an arbitrarily correlated two-component source with arbitrarily small probability of error over a capability-degraded broadcast channel (BC). A discrete memoryless BC $\{\mathsf{X}, P(y,z \mid x), \mathsf{Y} \times \mathsf{X}\}$ is said to be capability-degraded if $I(X;Z) \le I(X;Y)$ for all probability distributions $P(x)$ on X. These matching conditions are stated in a computable form. Secondly, we show that the sufficient conditions, obtained by Cover, El Gamal, and Salehi (1980) for reliable transmission of a two-component correlated source over a discrete memoryless multiple-access channel (MAC), are also necessary, if the source (S,T) satisfies the condition that S and T are conditionally independent given their common part K. The matching conditions obtained this way include several previously known cases for which necessary and sufficient conditions were found for sending a correlated source reliably over a MAC.

Suboptimal Link Scheduling in a Network of Directed Transceivers

Galen Sasaki *Department of Electrical and Computer Engineering, The University of Texas, Austin, TX 78712-1084*

The problem of scheduling data transfers in a network (V,E) of transceivers, where preemption of transfers is allowed and transmissions are directed, has been studied by Choi and Hakimi among others. In the network each node v has $b(v)$ transceivers, each link e can have at most $c(e)$ simultaneously communicating pairs of transceivers, and L_s is the size of the smallest odd cycle of (V,E) (hence, $L_s \ge 3$). Choi and Hakimi provided sub-optimal scheduling algorithms that have time complexity $O(|E|^2 \sum_{v \in V} b(v))$ and produce schedules of length at most $[1 + \frac{1}{L_s - 1}]\tau$ where τ is the length of the optimal schedule. For the case when $c([u,v]) \in \{0, 1\}$ (resp., $c([u,v]) \in \{0, \min\{b(v), b(v)\}\}$) for all links $[u,v]$, we provide an $O(|E|^2 |V|)$ (resp., $O(|E||V|^3)$) time algorithm that produces a schedule with length at most

$$\{1 + 2[(L_s + 1)\min_{v \in V} b(v) + L_s - 3]^{-1}\}\tau \quad (\text{resp., } \{1 + [L_s \min_{v \in V} b(v) - I[\min_{v \in V} b(v) \text{is odd}]]^{-1}\}\tau).$$

For the case when $c(e) \in \{0, 1\}$ for all links e, the scheduling algorithm can be modified to output the schedule in certain useful representations and have smaller time complexities ($O(|E||V|^2)$ or $O(|E|^2)$ depending on the representation).

feedback from the fusion center to the sensor platforms. In the appropriate probabilistic setting, it is well known that the minimum-mean-square-error estimate of X given $Y_1,...,Y_n$ is $\mathsf{E}[X \mid Y_1,...,Y_n]$. However, using any physical communication system, it is not possible to transmit real-valued quantities without distortion. Hence, it is not possible for the fusion center to compute $\mathsf{E}[X \mid Y_1,...,Y_n]$. In order to formulate a system model in such a way that we have some control over the distortion of the data transmitted to the fusion center, we shall consider the simultaneous optimization of the choice of a constrained fusion rule together with the choice of quantizers for the sensor platforms.

Recursive Pseudo Maximum-Likelihood Estimation for Joint Carrier Phase and Symbol Timing Recovery

Yih-Fu Won and Chung-Chin Lu *National Tsing-Hwa University, Institute of Electrical Engineering, Hsin-Chu, Taiwan, 30043 R.O.C.*

A recursive on-line algorithm for joint carrier phase and symbol timing recovery of linear-modulated systems, called recursive pseudo maximum-likelihood estimation (RPMLE), has been derived based on the maximization of a negative cost function consisting of pseudo likelihood observations. RPMLE uses the inverse of accumulated Hessian matrices of pseudo likelihood functions as weighting matrices to update timing parameters recursively. RPMLE can be applied to baseband PAM timing recovery by a similar approach. Several synchronizers are presented, which accommodate either data-aided or nondata-aided estimation according to whether detected data is directed into the estimators. Meanwhile, we find that the tracking loop implementation for joint carrier phase and symbol timing recovery developed by Meyers and Franks is a special case of RPMLE with constant weighting matrix. Simulation results show that the rate of convergence of RPMLE is faster than that of stochastic approximation. RPMLE owns high noise immunity and can be applied to solving synchronization problems of high transmission rate systems.

SESSION MP4

QUANTIZATION I

Adaptive Entropy-Coded Predictive Vector Quantization of Images

J. W. Modestino and Y. H. Kim *Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180* (40 min.)

We describe a new approach to image coding based upon adaptive entropy-coded 2-D predictive vector quantization (PVQ) ideas. PVQ is a straightforward *vector* extension of ordinary *scaler* predictive quantization schemes, such as DPCM, where the vector quantizer (VQ) is now embedded in the predictive feedback loop. Prediction is then performed on a vector, or block, basis using *previously* encoded blocks with the *prediction error* blocks subsequently applied, on a block-by-block basis, to the VQ. While PVQ is not new, previous applications have not attempted to exploit the further compressibility of the VQ output through use of variable-length entropy coding. In this paper we consider 2-D PVQ of images subject to an entropy constraint and demonstrate the substantial performance improvements over existing approaches. Furthermore, we describe a simple adaptive buffer-instrumented implementation of this 2-D entropy-coded PVQ scheme which can accommodate the associated variable-length entropy coding while *completely* eliminating buffer overflow/underflow problems at the expense of only a *slight* degradation in performance. This scheme, called 2-D PVQ/AECQ, is shown to result in excellent rate-distortion performance and impressive quality reconstructions on real-world images. Indeed, the real-world coding results shown here are rather striking and demonstrate almost imperceptible distortions at rates as low as 0.5 bits/pixel.

Necessary Conditions for the Optimality of Residual Vector Quantizers

Christopher F. Barnes and Richard L. Frost *Department of Electrical & Computer Engineering, Brigham Young University, Provo, Utah 84604*

Multistage or *residual* vector quantizers (RVQs) previously have been designed such that each stage satisfies conditions necessary for the optimality of single stage exhaustive search vector quantizers. Consequently, each stage of the RVQ is optimized independently of all other stages. This independent stagewise optimization causes performance to suffer whenever two or more stages are concatenated. In this paper the structure of residual quantizers is considered and the structural implications of these quantizers are made explicit for the RVQ encoder and decoder. Once the structural constraints are made explicit, we are able to determine necessary conditions for the *joint optimality* of all stages. We first determine necessary conditions for minimum mean square error for the quanta and partitions of residual *scalar* quantizers. The derivation of the scaler results motivate an approach yielding a vector generalization of the optimality conditions which hold for a broad class of distortion measures. An algorithm is given which yields RVQs which satisfy these necessary conditions and performance comparisons are given between conventionally designed RVQs and RVQs designed with the new algorithm.

Quantization for Decentralized Estimation from Correlated Data

M. Di Bisceglie and M. Longo *Selenia S.p.A., Radar Dept., Via Tiburtina, km. 12,400, I-00131 Roma, Italy, and Universita' di Napoli, Dipartimento di Ingegneria Elettronica, Via Claudio 21, I-80125 Napoli, Italy*

In a decentralized estimation scheme with two sensors data are remotely encoded by scalar quantizer to fulfill capacity constraints of channels conveying information to a central estimator. The quantizers are to operate separately, but a *joint* model of observations -- allowing for spatially correlated data -- may be taken into account at the design stage. We address the following problem.

Given: the source distribution P_X, the observation model P_{Y_1,Y_2}, a distortion measure $d(x, \hat{x})$; find quantizers $Q_1, Q_2 : \mathsf{Y}_1 \times \mathsf{Y}_2 \rightarrow \mathsf{M}_1 \times \mathsf{M}_2$ with rates $R_1 R_2$, and estimator $\hat{Z} = g(m_1 m_2)$, m_1, $m_2 \in \mathsf{M}_1 \times \mathsf{M}_2$; such that the average distortion $D = Ed(X, \hat{X})$ is minimum.

We prove that, for given Q_1 Q_2 and g, the sequence

1. $Q_1^+(y_1) = \text{argmin}_{m_1 \in \mathsf{M}_1} \int_X \sum_{m_2} d(x, \hat{x}(m_1, m_2)) p(x, m_2, y_1) dx$
2. $g^+(m_1, m_2) = \text{argmin}_{\hat{x}} \int_X d(x, \hat{x}) p(x \mid m_1, m_2) dx$

is a descent step for D. By alternately applying this iteration to both quantizers a design algorithm results.

For a linear Gaussian observation model we derive the limit performance of any encoding-decoding scheme, in terms of an upper bound to the rate-distortion function $G_G(D)$, where $R_G = R_1 + R_2$. Under various combinations of per-channel signal-to-noise ratios and spatial correlation coefficients, quantizers designed via the proposed algorithm nearly achieve such limit performance.

Optimal Quantization and Fusion in Multiple Sensor Systems with Correlated Observations

Yawgeng A. Chau and Evaggelos Geraniotis *Department of Electrical Engineering & Systems Research Center, University of Maryland, College Park, MD 20742*

In this paper we address two problems of quantization and data fusion from multiple sensors for the detection of a weak signal in dependent noise; (i) fusion from sensors with mutually independent observations and (ii) fusion from sensors with correlated observations. In the first problem, the observations of each sensor consist of a common weak signal disturbed by an additive stationary m-dependent, ϕ-mixing, or ρ-mixing noise process. The noise processes of the individual sensors are mutually independent. In the second problem the noise processes of the different sensors are correlated.

Three distinct schemes involving (a) fusing the test statistics formed by the sensors without previous quantization, (b) quantizing directly each of the sensor observations and then fusing, and (c) quantizing the test statistics of the sensors and then fusing them are considered. The memoryless nonlinearities, as well as the break-points and quantization levels of the quantizers introduced in these quantization/fusion schemes are obtained as solutions to appropriate integral equations which result from the optimization of suitable measures (error probabilities and deflection) of the performance of the fusion center.

Joint Vector Quantizer and Signal Constellation Design for the Gaussian Channel

Michael G. Perkins *The German Aerospace Research Establishment (DLR), NE-NT-T, Oberpfaffenhofen, 8031 Wessling/Obb., West Germany*

The usual approach to combined source/channel coding is to design the source and channel coders independently; however, in many cases significant gains can be achieved by jointly designing the two coders. This paper addresses the problem of joint vector quantizer and signal constellation design for the Gaussian channel. It is assumed that the transmitter is operating under a peak power constraint, that coherent modulation with perfect transmitter/receiver carrier-phase synchronization is employed, and that the receiver makes maximum aposteriori decisions. Under these assumptions, an algorithm for finding a locally optimal vector-quantizer/signal-constellation pair is presented.

The algorithm makes use of two sub-algorithms: one for optimizing the design of a vector quantizer for a given signal constellation, and one for optimizing the design of a signal constellation for a given vector quantizer. Starting with an initial vector quantizer and signal constellation, these two sub-algorithms are repeatedly applied, first one then the other, until neither the vector quantizer nor the signal constellation changes significantly from its previous state.

Simulation results for a low-rate image coder communicating over a noisy channel are presented.

Bennett's Integral for Vector Quantizers, and Applications

Sangsin Na and David L. Neuhoff *Dept. of Electrical Engineering, University of Nebraska - Lincoln, Lincoln, NE, 68588, and Department of EECS, The University of Michigan, Ann Arbor, MI 48109*

This paper extends Bennett's integral to vector quantizers and rth power distortion measures. The result is a simple formula that expresses the distortion of a many-point vector quantizer in terms of the source density, the number and density of quantization points, and the inertial density of the quantizer. The latter is a function that, at a given point $\mathbf{x}$, approximately equals the normalized moment of inertia of the quantization cells around $\mathbf{x}$. Previous extensions have not included the inertial density, which limited their applicability. The new version is formulated in terms of a sequence of quantizers whose point and inertial densities approach known functions. Precise conditions are given for the convergence of the distortion to Bennett's integral. A Bennett-like approximate formula for the entropy of the quantizer output is also developed.

Application of the new Bennett's integral leads to: a rigorous derivation of the distortion of multi-dimensional companders; the conclusion that for memoryless sources, the advantage of vector quantizers over scalar lies principally in their superior inertial densities, rather than point densities; and a quantitative analysis of the increase in distortion of various structured vector quantizers (e.g., tree-structured or block-transform) due to their formation of quantization cells with relatively few faces.

A New Algorithm for the Design of Locally Optimal Adaptive Vector Quantizers (AVQ)

G. Szekeres and G. Gabor *Information Theory Group, Hungarian Academy of Sciences, Budapest , Hungary, H-1111, and Dept. of Math. Stats. & Comp. Sci., Dalhousie University, Halifax, N.S. Canada*

After specifying an appropriate notion of optimality, three necessary conditions of optimality are established for forward adaptive VQ's, two of which are analogous to that of Lloyd and Gray *et al.* and the third is related to the adaptive nature of the VQ. Based on these conditions a convergent iterative algorithm is developed which uses either a long training sequence or the true distribution for the design of adaptivity (next-state function) and a set of VQ's with nonincreasing distortion in each step. The same algorithm is also used as a feed-forward design for backward adaptive VQ's. Experimental results with speech waveforms are also discussed.

Finite-State Vector Quantizers for Channel-Error-Resistance

Lei Ye and Zheng Hu *University of Electronic Science and Technology of China, Chengdu, P.R. China and Xidian University, Xian, P.R. of China*

Since a finite-state vector quantizer is a variety of tracking finite-state systems, channel errors can have disastrous effects. The degree of the effects and means to combat it are very important in theory and practice. In this paper, we study finite-state vector quantization of channel-error-resistance in noisy channels. The sufficient condition for error-finite propagation and the existence theorems for finite-state vector quantizers with bounded-error propagation are proved. A formula for estimating performance of error propagation in finite-state vector quantizers of channel-error-resistance is derived.

SESSION MP5

SHANNON THEORY I

A New Outlook on Shannon's Information Measures

Raymond W. Yeung *AT&T Bell Laboratories, Crawfords Corner, Holmdel, NJ 07733-1988*

We formalize the previous work of Reza, and Csiszar and Körner on the underlying mathematical structure of Shannon's Information measures. Let X_i, $i = 1, \ldots, n$ be discrete random variables. By regarding random variables as set variables, let $\Omega = \bigcup_i X_i$ be the universal set and F be the σ-field generated by $\{X_i\}$. We show that Shannon's information measures constitute a unique measure on F. To be precise, the Shannon information measure (i.e., Shannon's information measures as a whole) is a measure on F. This point of view, which we believe is of fundamental importance, has apparently been overlooked in the past by information theorists. As an immediate consequence we introduce the I-Diagram, which is a geometrical representation of the relationship among the information measures. The I-Diagram is similar to the Venn Diagram in set theory. We discuss the use of the I-diagram; some applications of which reveal previously unknown results. We also propose the mutual information measure for an arbitrary number of random variables and discuss some of its properties.

Finding a Basis for the Characteristic Ideal of an n-Dimensional Linear Recurring Sequence

Patrick Fitzpatrick and Graham Norton *University College, Cork, Ireland and University of Bristol, United Kingdom*

Let $\mathbb{N}$ denote the set $\{0, 1, \ldots\}$, and $\mathbb{N}^n$ the cartesian product of n copies of $\mathbb{N}$, and let $\mathbb{F}$ be a field. We denote the power series ring $\mathbb{F}[[X_1, \ldots, X_n]]$ by $\mathbb{F}[[\mathbf{X}]]$ and abbreviate the monomial $X_1^{i_1} \cdots X_n^{i_n}$ to $\mathbf{X}^{\mathbf{i}}$ for $\mathbf{i} \in \mathbb{N}^n$. Let $(\sigma) := (\sigma_{\mathbf{i}})$ be a sequence of elements from $\mathbb{F}$ indexed by $\mathbb{N}^n$.

If (σ) satisfies a *linear recurrence relation* of the form

$$\sum_{\mathbf{s} \in S} f_{\mathbf{s}} \sigma_{\mathbf{s}+\mathbf{i}} = 0 \text{ for all } \mathbf{i} \geq \mathbf{0}$$

where S is some finite non-empty subset of $\mathbb{N}^n$ and $f_{\mathbf{s}} \in \mathbb{F}$ for all $\mathbf{s}$, then (σ) is called an *n-dimensional linear recurring sequence* (or *$n-D$ lrs*) in $\mathbb{F}$. The corresponding polynomial

$$f(\mathbf{X}) := \sum_{\mathbf{s} \in S} f_{\mathbf{s}} \mathbf{X}^{\mathbf{s}}$$

in $\mathbb{F}[\mathbf{X}]$ is the *characteristic polynomial* of (σ) associated with the relation above. For convenience we define the zero polynomial to be a characteristic polynomial of every *$n-D$ lrs*.

It is easy to see that the set $\mathfrak{f}(\sigma)$ of characteristic polynomials of (σ) forms an ideal in $\mathbb{F}[\mathbf{X}]$: this is the *characteristic idea* of (σ). By Hilbert's Basis Theorem, $\mathfrak{f}(\sigma)$ has a finite set of generators. Our aim in this paper is to describe a constructive method (Algorithm IDEALBASE) by which such a basis may be determined, in the case that (σ) is *rectilinear*, that is, when $\mathfrak{f}(\sigma)$ contains a polynomial $p_k(X_k)$ with $p_k(0) \neq 0$ for each $k = 1, \ldots, n$. It will be observed that this assumption is reasonable, in view of the fact that our methods apply in particular to doubly periodic arrays, 2-dimensional cyclic (or TDC) codes, and also to more general polynomial codes in several variables. Moreover, our results may be applied to related areas such as the theory of the rational transfer functions associated with the synthesis of digital filters.

The Shannon-McMillan-Breiman Theorem and Other Information Theory Results Via a New Ergodic Theorem,

John C. Kieffer Department of Electrical Engineering, University of Minnesota, 200 Union Street, S.E., Minneapolis, Minnesota 55455

Let A be a finite set and let $X_1, X_2, \cdots$ be a stationary process with state space A. For each positive integer n, let F_n be a nonempty family of real-valued functions on A^n. A new ergodic theorem is presented, which, if certain assumptions about the sequence of families $\{F_n\}$ are satisfied, allows one to say the following: (1) there is an essentially unique function $f_n^* \in F_n$ for which $E f_n(X_1, \cdots, X_n)$ is minimized as f_n ranges over F_n; and (2) the random variable $f_n^*(X_1, \cdots, X_n)/n$ converges almost surely as $n \to \infty$. A sketch of the proof of this theorem is given. Also, a survey is given of some of the information theory results (including the Shannon-McMillan-Breiman Theorem) that are obtainable via the new ergodic theorem. (Research supported by NSF Grant NCR-8702176.)

Shannon's Coding Strategies for the Two-Way Channel--A Computer Attack

J. Pieter M. Schalkwijk *Department of Electrical Engineering, Eindhoven University of Technology, Den Dolech 2, P.O. Box 513, 560 MB, Eindhoven, The Netherlands*

Shannon derived an expression for the capacity region of the two-way channel (TWC) as the limiting rate, normalized with respect to the block length n, of optimum fixed length coding strategies of increasing block length n. These coding strategies can conveniently be represented as strategies for subdividing the unit square. We will report on the results of a computer study of these subdivisions of the unit square. The computer can yield effective strategies by simulated annealing. For very short ($n \leq 3$) strategies is it possible to find the best mixed strategy by exhaustively testing all possible sets of pure (i.e. Rows and columns of the unit square) strategies. The computer study is undertaken in a renewed effort to tightly upper bound the rate of certain deterministic TWC's, such as Blackwell's binary multiplying channel (BMC). We conclude by considering the prospects of achieving our ultimate goal.

A Sperner-Type Theorem and "Symmetric Versions" of Zero-Error Capacities

J. Körner and G. Simonyi *Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, P.O. Box 127, Hungary*

Several classical combinatorial problems arise through a certain symmetrization of the concept of zero-error capacity of compound discrete memoryless channels. We shall show that in some cases the more severe restrictions so imposed do not effect the asymptotic results. We prove in particular that if $N(n)$ denotes the largest cardinality of a family $\{D_1, D_2, \ldots, D_n\}$ of subsets of a set of n elements such that for every i the set D_i is the disjoint union of an A_i and B_i with

$$A_i \not\subset D_j,\ B_i \not\subset D_j \text{ for every } j \neq i,$$

then $[N(n)]^{1/n}$ converges to $(1+\sqrt{5})/2$.

Information Rates of Subsets and Matrix Inequalities

Thomas Cover and Joy Thomas *Departments of Electrical Engineering and Statistics, Stanford University*

The entropy per element of a randomly chosen subset of a set of random variables decreases with set size, thus proving Szasz's string of determinant inequalities. We show that the conditional entropy per element increases with set size, thereby proving a similar set of determinant inequalities. Yet another string of determinant inequalities follow from the monotonicity of mutual information between subsets of random variables.

Feedback in Discrete Communication

Alon Orlitsky *AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-2070* (40 min.)

$\mathcal{X}$ and $\mathcal{Y}$ are finite sets. (X,Y) is a random variable distributed over $\mathcal{X}\times\mathcal{Y}$ according to some probability distribution $p(x,y)$. Person P_x knows X, Person P_y knows Y, and both know p. They communicate in order for P_y to learn X. P_x may or may not learn Y. How many information bits must be transmitted (by both persons) in the worst case?

$C_1(p)$ is the number of bits required when only one message is allowed, necessarily from P_x to P_y. $C_2(p)$ is the number of bits required when only two messages are permitted: P_y transmits a message to P_x then P_x responds with a message to P_y. $C_\infty(p)$ is the number of bits required when P_x and P_y can communicate back and forth. Messages from P_y to P_x are called feedback.

The maximal reduction in communication achievable via feedback is one bit away from logarithmic: all probability distributions p have $C_\infty(p)\geq[\log C_1(p)]+1$ while there are distributions for which equality holds. therefore $C_1(p)$ can be exponentially larger than $C_\infty(p)$. Yet $C_2(p)$ cannot. With just one feedback message the number of bits required is at most four times larger than the minimal: $C_2(p)\leq 4C_\infty(p)+2$.

Surprisingly, for almost all sparse probability distributions, P_y who wants to say nothing must transmit almost all the bits in order to achieve the minimal number: $C_\infty(p)$. The number of bits transmitted by P_y can be appreciably reduced only if P_x transmits exponentially more than $C_\infty(p)$ bits.

If the communicators can tolerate ε probability of P_y not knowing the correct value of X and if randomized protocols are allowed then $4C_\infty(p)+2\log(1/\varepsilon)$ bits suffice even without feedback.

SESSION MP6

CODING THEORY II

Decoding is Really Hard

Jehoshua Bruck and Moni Naor *IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120-6099*

The problem of Maximum Likelihood Decoding of linear block codes is known to be hard. We show that the problem remains hard even if the code is known in advance and can be preprocessed for as long as desired in order to devise a decoding algorithm. The hardness is based on that an existence of a polynomial time algorithm implies that the polynomial hierarchy collapses. Namely, some linear block codes probably do not have an efficient decoder. The proof is based on results in complexity theory that relate uniform and nonuniform complexity classes.

A Generalization of the Discrete Fourier Transform in Finite Fields

Peter Mathys *Department of ECE, Box 425, University of Colorado, Boulder, CO 80309*

Let $\mathbf{v}$ denote a vector of length N over a finite field of characteristic p. Then a spectral representation $\mathbf{V}$ of $\mathbf{v}$ can be obtained by using the discrete Fourier transform (DFT) in a finite extension field of characteristic p, provided that p and N are relatively prime. It is shown that the DFT in finite fields can be generalized quite naturally to include block-lengths N which are divisible by p. The resulting spectral description of vectors over finite fields appears to be useful for the characterization of certain block codes. For the special case where N is a power of p, a simple characterization of (generalized) Reed-Muller codes can be obtained which is equivalent to the one given by Massey, Costello, and Justesen.

Upper and Lower Bounds on Aliasing Probability of Some Signature Analysis Registers

Toru Fujiwara, Feng Shou-ping, and Tadao Kasami *Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560 Japan*

In self-testing of LSI, a single or multiple input linear feedback shift register (LFSR) is widely used as a signature analysis register. An LFSR can be considered to be a circuit for dividing an input polynomial by a polynomial $g(x)$. It is known that the probability of an aliasing error in the LFSR is equal to that of an undetected error of a shortened cyclic code generated by $g(X)$ when the code is used for error detection. Here, we assume that the probability of an error occurring at an output bit of the circuit under test is a constant ε.

For a binary code C, let $P_e(C,\varepsilon)$ be the probability of undetected error of C for a binary symmetric channel with bit-error rate ε. Let HM_n^r denote a shortened $(n, n-r)$ Hamming code, and $RS_n^m(g(X))$ denote a binary code obtained from a shortened Reed-Solomon code over $GF(2^m)$ of length n (symbols) whose generator polynomial is $g(X)$. We present several upper and lower bounds on $P_e(C,\varepsilon)$ for HM_n^m and $RS_n^m(g(X))$ with $g(X) = (X-\alpha)$, $(X-1)(X-\alpha)$, $(X-\alpha)(X-\alpha^2)$, or $(X-1)(X-\alpha)(X-\alpha^2)$.

We show that

$$0.9\times 2^{-32} \le P_e(HM_n^{32}, 0.2) \le 1.1\times 2^{-32}, \text{ for } 303 \le n \le 2^{32}-1$$

and

$$0.9\times 2^{32} \le P_e(HM_n^{32}, 0.05) \le 1.1\times 2^{-32}, \text{ for } 1210 \le n \le 2^{32}-1.$$

We also show that

$$0.99\times 2^{-32} \le P_e(RS_n^{32}(X-\alpha), 0.2) \le 1.01\times 2^{-32}, \text{ for } 32 \le n \le 2^{32}-1,$$

and

$$0.9\times 2^{-32} \le P_e(RS_n^{32}(X-\alpha), 0.05) \le 1.1\times 2^{-32}, \text{ for } 63 \le n \le 2^{32}-1.$$

A Strengthening of the Assmus-Mattson Theorem

A. R. Calderbank, P. Delsarte, and N. J. A. Sloane *Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, NJ 07974; Philips Research Laboratories, Avenue van Becelaere, B-1170 Brussels, Belgium; and Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, NJ 07974*

Let $w_1 = d, w_2, ..., w_s$ be the weights of the non-zero codewords in a binary linear $[n, k, d]$ code C, and let $w_1', w_2', ..., w_{s'}'$, be the non-zero weights in the dual code $C^{\perp}$. Let t be an integer in the range $0 < t < d$ such that there are at most $d - t$ weights w_i' with $0 < w_i' \le n - t$. Assmus and Mattson proved that the words of any weight w_i in C form a t-design. Let $\delta = 0$ or 1, according as C is even ($w_{s'}' = n$) or not, and let B denote the set of codewords of weight d. We prove that if $w_2 \ge d + 4$, then either (1) $t = 1$, d is odd, and B partitions $\{1, 2, ..., n\}$, or (2) B is a $(t + \delta + 1)$-design, or (3) B is a $\{1, ..., t + \delta, t + \delta + 2\}$-design. If C is a self-orthogonal binary code with all weights divisible by 4 then the result extends to codewords of any given weight. The special case of codewords of minimal weight in extremal self-dual codes also follows from a theorem of Venkov and Koch; however our proof avoids the use of modular forms.

Reduced Lists of Patterns for Maximum Likelihood Soft Decoding

Jakov Snyders *Department of Electronic Communications, Control and Computer Systems, Tel-Aviv University, Ramat-Aviv 69978, Israel*

Maximum likelihood soft decision decoding of an (n,k,d) binary linear block code is performable by complementing the hard-detected version of the received word in at most $m = n - k$ positions. The set of positions, or pattern, is selected according to least sum of reliabilities of the associated bits. We introduce a method whereby out of all the patterns with cardinality m, $m-1$ and $m-2$ no more than, respectively,

$$1,\ \delta_{d4} + \binom{m-d+3}{2} \text{ and } \left[\binom{m}{3} - 3\binom{m-1}{3}\right] \delta_{d3} + \binom{m-1}{3}(\delta_{d4} + \delta_{d5}) + 6\sum_{i=d}^{m-1}\binom{m-1}{i}$$

explicitly described patterns have to be scored, provided that $d \ge 3$. Here δ is the Kronecker delta. This approach enables decoding of the (15,11,3) code by at most 51 real additions, compared to the previous best 83 additions required in the worst case by a different search scheme. Decoding of the (31,26,3) and (32,26,4) codes is accomplished by less than 200 additions in the worst case, versus more than 1000 additions performed by any previously published decoder. Application of reduced lists of patterns to coset-decoding of medium rate codes is also addressed.

Bounds on Codes via Kolmogorov Complexity

John T. Coffey and Rodney M. Goodman *California Institute of Technology, Pasadena, CA 91125*

In this paper, we examine the use of the theory of Kolmogorov complexity in analysis of the error-correction capabilities of various classes of codes (e.g., systematic linear codes, shortened cyclic codes, generalized Reed-Solomon codes) under various error-correction strategies (random-error correction, burst-error correction, and various mixed strategies). Many original results are given; in addition, the techniques used are new and intuitively appealing. Some comments are added about the problem of giving explicit constructions for codes that meet these bounds, and we also give one additional useful result derivable from the principles of Kolmogorov complexity.

Bounds on the Dimension of Certain Codes and Subcodes

Alexander Vardy, Jakov Snyders, and Yair Be'ery *Department of Electronic Communications Control and Computer Systems, Tel-Aviv University, Ramat Aviv 69978, Tel-Aviv, Israel*

Let C be an (n,k) linear block code over $GF(q)$, generated by a matrix G. A nonnegative integer λ is said to be the *contraction index* of C if a maximal set of pairwise linearly independent columns of G has $k+\lambda$ elements. In recent works of Conway & Sloane, Be'ery & Snyders and Forney it was shown that the complexity of soft decision decoding of block and lattice codes can be considerably reduced by means of partitioning the code into cosets with respect to a subcode of large dimension and small contraction index. In this paper we derive several upper and lower bounds on the dimension of a subcode with a prescribed contraction index. We also present an upper bound on the dimension of any (n,k) code over $GF(q)$, with minimum Hamming distance d and contraction index λ. For certain values of n and d, the latter bound is shown to be tight for all q and λ. This substantially generalizes the results obtained previously for $\lambda = 1$ in the context of majority logic decoding.

Achieving the Cutoff Rate on Communications Channels

J. T. Aslanis and J. M. Cioffi *Information Systems Laboratory, Stanford University, California*

In this paper we claim that combining the best current coding techniques with appropriate equalization can achieve the cutoff rate on channels with memory at any SNR. We shall examine characteristics of codes, equalizers, and signal sets that both the cutoff rate and capacity formulas indicate are necessary for this performance. The capacity and cutoff rate formulas indicate the optimal bandwidth and bit spectral efficiency (bits/symbol/Hz) for a given input signal power constraint. We show that the infinite blocklength limiting performance of vector coding, an approximation to a multiple carrier technique combined with trellis coding, nearly achieves the cutoff rate. We then bound the difference between capacity with and without power spectral shaping and note that this difference, which equals the vector coding power penalty, becomes important only at very low SNR. In this range of SNR, below one bit/symbol, trellis codes cannot provide adequate coding gain, but convolutional codes (or block codes) can easily be combined with the vector coding "equalization" techniques to achieve cutoff rate performance, with moderate complexity. We show that, for a fixed SNR, the capacity and cutoff rate R_0^* differ by less than 1-1/2ln2 bits per symbol on any channel with or without ISL. We finally note that, while the full 1.53 dB of shaping gain affects the cutoff rate at high SNR, the total gain cannot be obtained if one fixes the available bits/dimension in a multi-dimensional signal set.

SESSION MP7

ERROR-CONTROL AND OTHER CODING

The Rate/Performance Tradeoffs of Focused Error Control Codes

Tom Fuja and Fady Alajaji *Department of Electrical Engineering, Systems Research Center, University of Maryland, College Park, MD 20742*

Let B be a set of non-zero elements of $F_q(q>2)$; we say a code is (t_1,t_2)-focused on B if it can correct up to t_1+t_2 errors *provided* at most t_1 of those errors lie outside B. The strategy is to offer different levels of protection against "common" errors - those in B - and "uncommon" errors. (The motivating example: correction of a single-bit-per-byte errors with codes over F_{2^8}.)

This talk will compare the performance and rates of (t_1,t_2)-focused codes with those of traditional t_1+t_2-error correcting codes. We show that, at high SNR, if a channel is sufficiently "skewed" - that is, if the noise character is Z and $P\{Z \notin B \mid Z \neq 0\} < \gamma_{crit}$ - then the performance of a (t_1,t_2)-focused code is essentially identical to that of a t_1+t_2- error correcting code; this claim is derived analytically and verified by simulation results. Since (t_1,t_2)-focused codes can be constructed with higher rates than can t_1+t_2-error correcting codes, they offer for these "skewed" channels new advantages in terms of rate and/or performance. We include in the talk an analysis of the tradeoffs offered by focused codes for M-ary PSK and M-ary ASK modulation schemes.

Design and Implementation of Binary Combined Error Control and Line Coding: a BCH-based Example

J. J. O'Reilly, S. Williams, and A. Popplewell *School of Electronic Engineering Science, University of Wales, Bangor, Dean Street, Bangor, Gwynedd LL57 1UT, United Kingdom*

There has been considerable interest expressed recently in defining transmission codes, which combine error control and line coding features both for telecommunications transmission and magnetic recording. This paper introduces the design of such a novel combined error control and line coding scheme, and describes its successful practical realization using a BCH example. The design proceeds by constructing linecodes from known error control codes and in this way we are able to form a linecode which is constructed entirely of valid error control codewords which thus inherits its error control power. The design is applicable to a wide range of error control codes, but attention is focused on its application to standard BCH codes when considering practical implementation. The implemented system is functionally described and the power spectra consequently produced by its encoded output are presented and compared with those derived by theoretical analysis. It is seen that these exhibit the characteristic DC-null of a conventional linecode.

The Design of Error Control Codes for Rayleigh Fading Channels with Memory

Jean-Claude Belfiore *Ecole Nationale Supérieure des Télécommunications, 46, rue Barrault, 75634 Paris Cedex 13, France*

The aim of this paper is to give a method to design codes adapted to the Rayleigh fading channel with memory. We first derive a Chernoff bound of the bit error probability after decoding. This bound shows that coding yields an equivalent order of diversity which increases with the average signal-to-noise ratio until a limit value which is the minimum (or free) distance of the code. We show that, in order to minimize the BER, the transmitted symbols must have unequal energy. The optimal energy distribution has been computed for some short block codes. It provides important gains over the uniform one.

Error-Control Coding for the Binary N-user Modulo-q Channel

Vivek Telang and Mark Herro *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556*

Efficient use of transmission facilities often requires the sharing of resources by a number of users. Commonly used methods of sharing resources are Time Division Multiplexing (TDM) and Frequency Division Multiplexing (FDM). In this paper, we describe a different approach to designing an N-user Multiple-Access System for the Binary Modulo-q Channel. We design a coding scheme for N users, such that when all N users are sending data, the channel is used at its maximum capacity. However, when fewer than N users are active, the unused channel capacity is used to achieve error control. The extent of error control is inversely proportional to T, the number of active users. The multiple-access codes described in this paper are based on well-known error-correcting codes, *viz.*, BCH codes, Reed-solomon codes and convolutional codes. Other codes that are being currently studied are cyclic codes and some Unequal Error Protection (UEP) codes. Possible applications of these codes are in a multiple-access system used by data sources as well as voice sources. The codes exploit the 50% activity of a typical bursty voice source to afford error-protection for the data sources without requiring additional channel bandwidth.

On the Decoder Error Probability of Linear Codes

Kar-Ming Cheung *Communication Systems Research Section, Jet Propulsion Laboratory, 4800 Oak Grove Dr., Pasadena, CA 91109*

In a recent paper by the author, an explicit formula which enumerates the complete weight distribution of an (n,k) linear block code using a partially known weight distribution is derived. Also an approximate formula for the weight distribution of most linear block codes is given, and is shown to be approximately binomial.

In this paper, by generalizing the coding and combinatoric techniques mentioned in the above paper, an approximate formula for the weight distribution of decodable words of most linear block codes is evaluated. This formula is then used to give an approximate expression for the decoder error probability $P_E(u)$ of linear codes, given that an error pattern of weight u has occurred. It is shown that $P_E(u)$ approaches the constant Q as u gets large, where Q is the probability that a completely random error pattern will cause decoder error.

Construction of Linear Codes of Minimum Distance Five

C. L. Chen *IBM Corporation, P.O. Box 950, Poughkeepsie, NY 12602*

In this paper, we present a method of constructing binary linear codes that have a minimum Hamming distance of five. Some of the new codes obtained are more efficient in information rates than other known codes.

The Permutation Channel

Jonas Wallberg and Ingemar Ingemarsson *Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden*

Permutation Modulation, PM, is a seldom used modulation scheme. It has however many favorable features like: each codeword requires the same energy, an easily implementable maximum likelihood receiver, the probability of an incorrect detection is the same for each sent signal.

We will here describe a special abstract channel which we call the Permutation Channel. The Permutation Channel contains the Additive White Gaussian Noise Channel. The vector model of the AWGN Channel is used. The set of transmitted vectors is obtained by permuting the components of a given initial vector in all possible ways. The PM detector maps the received vector onto the set of transmitted vectors.

For this channel we have computed upper bounds on the probability for different kinds of permutation errors. These error probabilities are important when constructing codes for the Permutation Channel. The performance of Permutation Modulation and Permutation Codes is compared to the performance of ordinary block codes with amplitude and phase modulation.

A Class of Error Correcting Codes for the Permutation Channel

Jonas Wallberg and Ingemar Ingemarsson *Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden*

The permutation channel is a discrete, memoryless channel which permutes the order of the transmitted components of a vector (see contribution "The Permutation Channel"). A *Permutation Code* is a subset C of the set of vectors Ω (in the Euclidean space) obtained by Permutation Modulation, *variant 1*. Thus a codeword is described by a permutation, p, of the components of the initial vector s_0. The error correcting codes will be

SHANNON LECTURE

Tuesday, 8 - 8:50 a.m.

A Factor of 2

Thomas Cover, *Stanford University, Stanford, California 94305*

TECHNICAL SESSIONS

Tuesday, 9 a.m. - 12 m.

SESSION TA1

COMMUNICATION THEORY I

A Model for the Statistical Analysis of Sigma-Delta Modulation

Ping Wah Wong and Robert M. Gray *Dept. of Electrical and Computer Engineering, Clarkson University, Potsdam, NY, 13676, and Information Systems Laboratory, Dept. of Electrical Engineering, Stanford University, Stanford, CA, 94305*

A stochastic model is suggested that describes the behavior of a single-loop sigma-delta modulator, which is an oversampled feedback quantization scheme. When the input to this nonlinear system is an i.i.d. Gaussian process, the model can be represented by a Wiener process embedded in a renewal process. A condition is given so that the difference between the output predicted by the model and that of the true system can be made arbitrarily small. Using this model, we can show the convergence and mixing properties of the output of the single-loop sigma-delta modulator. The power spectral density of the output sequence can also be derived. (Research supported by NSF Grant MIP87-06539 and a Stanford University Center for Integrated Systems seed grant.)

Error Probability for Digital Transmission over Nonlinear Channels with Application to TCM

Yow-Jong Liu, Ikuo Oka, and Ezio Biglieri *Electrical Engineering Department, University of California, Los Angeles, CA 90024-1594*

We consider the computation of error probability for digital transmission over nonlinear channels with a finite memory. It is well known that trellis-code modulation (TCM) encoders can be modeled as nonlinear systems with finite memory. To prove that the converse is also true, we use orthogonal Volterra series to derive a "canonical" representation for discrete nonlinear systems, based on a linear convolutional code and a memoryless mapper. This representation shows that finite-memory discrete nonlinear systems can be analyzed in much the same way as TCM schemes. In particular, TCM over nonlinear channels can be analyzed.

The "pairwise state" technique is in principle applicable to the computation of error probabilities for nonlinear channels. It involves $N^2 \times N^2$ matrices, where N is the number of states in the trellis representation of the overall channel. If TCM over a nonlinear channel is considered, N is the product of the number of code states and the number of channel states, and this may be very large.

We derive an upper bound to the error probability that avoids the consideration of $N^2 \times N^2$ matrices. It is based on the computation of the transfer function of a graph with $N+1$ nodes, whose branch labels are *matrices* rather than scalars. As a result, consideration of situations involving a relatively large number of states is made possible. A lower bound to the error probability is also derived. Examples of application to the analysis of a number of situations involving nonlinear channels are also provided.

Probability Distribution of DPSK in Tone Interference and Applications to SFH/DPSK

Q. Wang, T. A. Gulliver, and V. K. Bhargava *Department of Electrical and Computer Engineering, University of Victoria, P.O. Box 1700, Victoria, B.C., Canada V8W Q2Y2*

Slow frequency hopped differential PSK (SFH/DPSK) has been the subject of much recent attention. The most salient feature of SFH/DPSK is that it can sustain a much higher data rate than a fast frequency hopped system with the same hop rate. We present a study of the probability distribution of the

received DPSK signal under tone jamming to facilitate the analysis of a SFH/DPSK system.

The results given are more general than those previously published in several aspects. First, the differential phase of the transmitted DPSK signal can assume any value. Second, probability distributions are derived instead of a set of probabilities calculated over certain symmetrical regions. This allows maximum flexibility in performance analysis as far as decision regions in demodulating DPSK are concerned. Third, the joint probability distribution of both the magnitude and differential phase of the jammed DPSK signal is given. This can be used in the analysis when both tone jamming and Gaussian noise are considered. To illustrate the application of these results, we analyze the error probability performance of a general uncoded SFH/DPSK signal under worst case tone jamming and Gaussian noise.

Estimation Variance Bounds of Importance Sampling Simulations in Digital Communication Systems

D. Lu and K. Yao *Electrical Engineering Department, University of California, Los Angeles, Los Angeles, CA 90024-1594*

In practical applications of importance sampling (IS) simulation, we encounter two basic problems, that of determining the estimation variance and evaluating the proper IS parameters needed in the simulations. We derive a simple upper bound on the estimation variance which is applicable to all known importance sampling techniques. Furthermore, a lower bound on the improvement ratio of various importance sampling technique relative to the direct Monte Carlo simulation is also given. These bounds are shown to be useful and computationally simple to obtain. Based on these bounds, we can readily find the needed sub-optimum IS parameters. Specific numerical results indicate that these bounding techniques are applicable to many problems involving Gaussian and non-Gaussian pdf's, linear as well as non-linear communication systems with ISI in which exact bit error rates and IS estimation variance cannot be obtained simply by previously known approaches.

Bit Error Simulation via Conditional Importance Sampling

Tao Chen and Charles L. Weber *Communication Sciences Institute, Department of Electrical Engineering-Systems, University of Southern California*

A new variation of *importance sampling*, designated as *conditional importance sampling*, or CIS, is proposed for the simulation of bit error rate in nonlinear digital communication systems. *Monte Carlo* simulations are simple and tractable in simulating the bit error rate, but is prohibitively slow when errors are rare. *Importance sampling* is a variation of *Monte Carlo* which increases the simulation speed by altering the input density functions and weighting the output to have an unbiased estimator. Earlier research concentrated on a fixed amount of bias in the input noise probability density function, or pdf, for every simulation trial. We extend the biasing to the phase error pdf, and, in order to utilize more of the knowledge we have about the system, we suggest *conditional importance sampling*. CIS biases the pdf of some input variables according to the values of the other ones. The estimator variances of CIS are evaluated for some simple systems, and the resulting improvements over standard *Monte Carlo* are significant.

A Contribution to the Proof of the Simplex Conjecture

Dejan E. Lazić *Universität Karlsruhe, Institut für Algorithmen und Kognitive Systeme, D-7500, Karlsruhe 1, Fed. Rep. of Germany*

In order to prove the simplex conjecture (the set of $M=N+1$ unit vectors in N-dimensional Euclidean space that form a regular simplex is the optimum error protecting code for memoryless-channel with discrete time and additive Gaussian noise), two lemmas are established in this paper. The first one purports that the probability of error for a Voronoi cell (a radial projection of the maximum likelihood region to the surface of a unit N-dimensional sphere Ω_N) that forms a regular spherical simplex on Ω_N, that has N sides, and whose center of inscribed spherical cap contains a corresponding code word (this

cell will be called optimal Voronoi cell), is always less than the probability of error for another Voronoi cell of the same area as the optimal one, regardless of the position of its corresponding code word in its interior. The second lemma asserts that the average probability of error for optimal Voronoi cells is the least when all their areas are equal, given that the total area is constant. Lemma 1 is proved for $N=3$ and $N=4$, and lemma 2 for every $N>1$, establishing the validity of the simplex conjecture for $N=4$ (till now the conjecture was proved for $N<4$).

Joint Synchronization and Detection from Multiple Samples per Symbol

Costas N. Georghiades and Marc Moeneclaey *Electrical Engineering Department, Texas A&M University, College Station, TX 77843, and Communications Engineering Lab., University of Ghent, Belgium*

We investigate the problem of sequence estimation in the presence of a timing uncertainty from samples of the output of a matched filter taken at an integer multiple of the symbol rate. The effects of sampling rate and signaling pulse shape on both delay and sequence estimation performance are studied by means of lower bounds on mean-square estimation error, and error-probability respectively. Further, algorithms for delay and sequence estimation, based on joint maximum-likelihood processing of the data samples, and an iterative algorithm motivated by the expectation-maximization (EM) algorithm are introduced. The general results are applied to the case of binary antipodal signaling and rectangular pulse shapes.

Performance Analysis of the Sequential Algorithm for Intersymbol Interference Channels

F. Xiong and E. Shwedyk *Department of Electrical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2*

A receiver consisting of a whitened matched filter (WMF) followed by a sequential algorithm (SA) has been developed for intersymbol interference (ISI) channels. The channel-WMF system can be modeled as a finite state machine (FSM). The present work gives performance analysis. The symbol error probability is given by $KQ(\frac{d_{\min}}{2\sigma})$, where constant K depends on the channel; σ is the spectral density of the Gaussian noise; $d_{\min}$ is the minimum distance between received data sequences. To determine the computational complexity, the FSM can be interpreted as a special convolutional encoder followed by a binary symbol to Q-ary symbol mapping ($Q = 2^L$, L is the length of the channel impulse response). It follows that the computational distribution is Pareto and there exists a computational cutoff rate R_{comp}. For the uncoded data considered the rate is fixed and the R_{comp} criterion translates into a signal to noise (SNR) criterion, i.e., for bounded computation per node the SNR should exceed SNR_{comp}. An upper bound SNR'_{comp} on SNR_{comp} is found analytically by assuming a uniform input distribution. Computer simulation results verify the above analysis and show that for a one-pole channel the SA is 1-3 dB better than the M-algorithm.

Minimum Error Probability for Asynchronous Multiple Access Uncorrelated Fading Intersymbol Interference Channels

Daoben Li *Dept. of Information Theory, Beijing University of Posts and Telecommunications, Beijing 100088, China*

Consider an L-user multiple access digital communication system,in which all the subchannels corrupted by additive white Gaussian noise are uncorrelated fading intersymbol interference channels with the same statistical properties, and the users transmit independent data streams by modulating antipodally a set of assigned signal waveforms without maintaining relative bit-synchronism among them. The minimum error probability achieved by the maximum-likelihood sequence detectors is studied first. The corresponding optimum signal set and their physical realizability are next studied. It is shown that if the inphase and quadrature components of the signals are complementarily like a group of

bandlimited white noises, there is no performance degradation due to the presence of other users, and the intersymbol interference is a beneficial factor. The wider the system bandwidth or the larger the time dispersion of the channels, the better the system performance. In the limiting case the multiple-access data-transmission system with intersymbol-interference channels would approach that of a Gaussian single-use channel.

SESSION TA2

MODULATION

A Coded Modulation Scheme with Interblock Memory

Mao-chao Lin *Department of Electrical Engineering, National Taiwan University, Taipei 10764, Taiwan, ROC*

We present a block coded modulation scheme for which the coded modulation of each block depends on the preceding block. The design is a combination and modification of Sayegh's coded modulation scheme and a Sloane-Reddy-Chen's binary code construction. Using signal set geometry of 8-PSK and 8-AMPM, we design coded modulation systems with code rates of 29/45 and 15/24 respectively. The asymptotically coding gains are 6.02 dB and 5.05 dB compared to the uncoded 4-PSK, for which the decoding complexity are respectively similar to those for the 32-state and 16-state trellis coded modulations designed by Ungerboeck. If we consider the effect of the number of nearest neighbors, the coding gains are 5.34 dB and 4.52 dB respectively.

Combined Coding and Modulation Using Block Codes

Klaus Huber *Institut für Netzwerk- und Signaltheorie, Technische Hochschule Darmstadt, Merckstr. 25, 6100 Darmstadt, West-Germany*

In this contribution we present combined coding and modulation schemes which are simple, powerful, and easy to synchronize. At the same data rate and power consumption they improve considerably over uncoded signaling schemes. The essential new idea is the contraction of a n_0 bit-sequence to a $(n_0 - l)$ signal point-sequence, giving a redundancy of l bits at our disposal, which can be used for example for error control coding. An important advantage is, that the schemes are easy to implement. Also the problems of synchronization are much better to overcome.

Modulo Sigma-Delta Modulation for a Random Process Input

Wu Chou and Robert M. Gray *Information Systems Laboratory, Department of Electrical Engineering, Stanford, CA 94305*

Modulo sigma-delta modulation (MSDM) has its important application in communication and oversampled analog-to-digital conversion. MSDM is formed by having a modulo-limiter as the front end to compress the input into the non-overload region of the sigma-delta modulator. An exact analysis of the response of this nonlinear system with regard to an input with independent increments has been obtained.

It is shown that the normalized binary quantization noise process $\hat{\varepsilon}_n$ of MSDM, when the input is a process with independent increments, is a quasi-stationary process. It is uniformly distributed in $[-1/2, 1/2]$. It is asymptotically white and uncorrelated with the input. This is a significant difference from the case of dithering where the binary quantization noise is never white.

The method we use here provides an effective way to decide the asymptotic distribution of the quantization noise and its spectrum. It can also be applied to other modulation schemes such as modulo-PCM, modulo-DPCM and modulo multi-stage sigma-delta modulation.

Embedded Modulation and Coding for HF Channels

B. Honary and M. Darnell *Department of Engineering, University of Warwick, Coventry, CV4 7AL, UK and Department of Electronic Eng., University of Hull, Hull, HU6 7RX, UK*

In the proposed paper a number of new concepts and techniques intended for application to radio systems operating over time variable channels will be discussed. These are currently being investigated

via several co-ordinated research programs by the authors, with the authors, with the aim of integrating them progressively into the architectures of adaptive HF radio systems. However, it should be stressed that the techniques are not specific to HF radio: in principle, they could also be applied to other forms of radio system with non-Gaussian noise and interference backgrounds.

The techniques, which are studied and simulated in software form are implemented using Digital Signal Processor) (i) a new method for implementation of embedded codes, (ii) embedded modulation, and (iii) real-time channel evaluation obtained by varying constellation angles.

Design of (0,1) Sequence Sets for Pulsed Coded Systems

F. Khansefid, H. Taylor, and R. Gagliardi *Communication Sciences Institute, University of Southern California, Los Angeles, CA 90089*

A (0,1) sequence is a sequence of zero (off) and one (on) binary symbols of a given length. A set of sequences is a group of such (0,1) distinct sequences. The sequences can be considered as pulse waveforms, in which the one symbol corresponds to the location of an electronic or optical pulse while a zero prepresents a zero, or pulse absence. The design of (0,1) sequence sets therefore correspond to the design of families of pulsed waveforms that can be used as a codeset, and therefore has application to on-off pulse signaling in communication and signal processing systems.

This paper presents the results of a study relating the above sequence parameters to the ability to produce good autocorrelation and cross-correlation properties. Sequence sets having the optimal correlation properties are derived and tabulated for various parameters combinations. Best possible 4 - pulse sequences with pairwise correlation ≤ 1 have been found where the numbers of sequences in the set is any number up to 14. For 5 - pulse and 6 - pulse sequences, results close to the Kløve bound have been derived. A general construction is given for such sets, and the results of a computer search for optimal code sets is reported. Design bounds on the set parameters are also presented.

VLSI Viterbi Decoder for a BCM Code

Roksana Boreli, David Coggins, Branka Vucetic, and Shu Lin *Laboratory for Comm. Sci. and Engineering, Sydney University Electrical Engineering, N.S.W. 2006 Australia, and University of Hawaii, Manoa, HI, 96822*

This paper analyzes the VLSI implementation of a high-speed Viterbi decoder for a 4-state Block Coded Modulation (BCM) code with 8-PSK symbols. It demonstrates that practical decoders of low complexity for TDMA applications can be built on one chip for 100*M Bps* data rates. The device is being implemented in $1.5\mu m$ CMOS technology. The problems of synchronization, quantization and metric calculation are investigated. A comparison in terms of speed and decoder complexity is made on the register exchange and traceback methods, and it is shown that, for this particular code, the former is more suitable.

On the Capacity of the Gaussian Channel with a Finite Number of Input Levels

L. H. Ozarow and A. D. Wyner *AT&T Bell Laboratories, Murray Hill, NJ 07974*

We give an analytic confirmation of an observation made by Ungerboeck (in "Channel Coding with Multilevel/Phase Signals," *IEEE Trans. on Information Theory*, Vol. IT-28, Jan. 1982, pp. 55-67) that approximately optimal performance on a Gaussian channel with capacity C can be obtained with about 2^C levels. In particular our results imply that by using about 2^{C-1} levels, we can achieve a rate of nearly $C-1$ bit, and that by using about 2^{C+1} levels we can achieve a rate of about C -0.4 bits.

Enumerative Coding for Constrained Noiseless Channels and Modulation Coding

Boris Fitingof *Optical Sciences Center, University of Arizona, Tucson, AZ 85721*

Practically any group of constraints defining requirements for modulation codes can be described by a state transition table (STT) with finite number s of states (e.g., for any run-length-limited (d,k) code $s = k + 1$). Synchronization words can be easily introduced as additional constraints by modifying STT. New coding-decoding algorithms for constrained noiseless channels are proposed. These algorithms have been implemented in *C* under UNIX. For any given STT and any given length L of codewords the algorithms satisfy modulation constraints defined by this STT and provide optimal rate under these modulation constraints and for given L. Hence the rate converges to the Shannon's capacity for the noiseless channel when $L \to \infty$. For the case of the run-length-limited codes computer experiments show rapid convergence of the rate to the capacity with increasing input block length. The proposed algorithms, for both coding and decoding, (1) use as an additional input an array of $s \times L$ integers completely determined by the STT and L, (2) practically do not use other memory besides this array, (3) use the number of operations of order s per bit of encoded data, (4) achieve aforementioned results by enumeration of all output words satisfying the modulation constraints, using the enumeration algorithm for decoding and an algorithm inverse to the enumeration algorithm for coding.

SESSION TA3

OPTICAL COMMUNICATIONS

Trellis Codes for the Optical Direct-Detection Channel

Gregory J. Pottie *Codex Corporation, 20 Cabot Blvd., Mansfield, MA 02048*

It is shown that the reduced complexity minimum distance search method of Zehavi and Wolf can be extended from codes designed using the squared Euclidean distance measure to any measure which is additive over the code branches. Recently Georghiades has derived a distance metric for codes for the direct detection optical channel. The method is applied to find codes for four and eight point overlapping pulse position modulation (OPPM), with large minimum distance and a small number of codewords at that distance.

Application of Quantum Minimax Rule to General Ternary Quantum State Signals

Masahiko Sekiguchi, Osamu Hirota, and Masao Nakagawa *Faculty of Science and Technology, Keio University, 3-14-1, Hiyoshi, Hohoku-ku, Yokohama 223, Japan, and Faculty of Technology, Tamagawa University, 6-1-1, Machida, Tokyo 194, Japan*

A receiver which minimizes an effect of quantum noise is called quantum optimum receiver in optical communications. So far, the quantum Bayes rule and Neyman-Pearson rule were formulated by Helstrom. These rules are attractive, but very difficult to apply to arbitrary quantum state signals, except for binary signals and a few specific ones. On the other hand, the quantum minimax rule was formulated by Hirota, which provides a simplification of the calculation for the quantum detection problems. However, it seems to us that his method is still complex in order to solve the problems for arbitrary quantum signals.

In this presentation, a new calculation method of the quantum minimax rule for the pure state signals is proposed, employing a geometrical interpretation. The calculation process of N^2 equations obtained from the necessary and sufficient conditions of the minimax rule is greatly simplified by our method. As an example, the general solution for ternary signals is given. Then it is numerically shown that the optimum receiver based on the minimax rule is always superior to the quasi-classical receiver in case of the general ternary signals.

Error Probabilities in Optical Receivers with Avalanche Diodes

Chia Lu Ho and Carl W. Helstrom *Dept. of Electrical & Computer Engrg., R-007, University of California, San Diego; La Jolla, Calif. 92093.*

Error probabilities in optical receivers using avalanche diodes are calculated by numerical integration of Laplace inversion integrals in the complex plane along paths passing through a saddlepoint. (The communication system is transmitting information coded into equally probable 0's and 1's, represented respectively by a blank and by the transmission of a light pulse along a fiber). The integrand involves the moment-generating function of the amplified output of the diode, and this is calculated from Personick's model of the avalanche multiplication process. We take into account the shape of the incoming light intensity pulses, including intersymbol interference, and consider both exponential and Gaussian current pulses at the amplifier output.

Error Probability of Optical Feedback Receivers

Göran Einarsson *Telecommunication Theory, Lund University, Box 118, S-221 00 Lund, Sweden*

An analysis of noncoherent optical feedback receivers is presented. The signal statistics at the decision point are derived and the error probability is determined.

The negative feedback generates a shot noise process which is the difference between two Poisson point processes. The probability distribution of such a process is studied and compared with Gaussian and Poisson distributions.

The error probability is calculated for an ideal receiver model consisting of a feedback amplifier with frequency-independent gain followed by an integrate-and-dump filter. A comparison is made between the exact error probability and the results obtained from a Gaussian approximation. It is shown that the straight-forward Gaussian approximation works well and gives results close to the exact values.

A more realistic receiver model with arbitrary filter is studied. A saddlepoint approximation is derived which can be used to calculate the error probability with high accuracy.

Electrical Signal Processing Techniques in Long-Haul, Fiber-Optic Systems

Jack H. Winters and Richard D. Gitlin *AT&T Bell Laboratories, Holmdel, New Jersey 07733*

The purpose of this paper is to demonstrate the potential for electrical signal processing to mitigate the effect of intersymbol interference in long-haul, fiber-optic systems. Intersymbol interference in long-haul fiber-optic systems can severely degrade performance and consequently limit both the maximum distance and data rate. The sources of intersymbol interference include nonlinearity in the transmit laser, chromatic dispersion in systems operated at wavelengths other than the dispersion minimum of the fiber, polarization dispersion, and bandwidth limitations in the receiver. We discuss several techniques for reducing intersymbol interference in single-mode fiber systems with single-frequency lasers and show which techniques are appropriate at high data rates in direct and coherent detection systems. In particular, we analyze the performance of linear equalization (tapped delay lines), nonlinear cancellation (variable threshold detection), maximum likelihood detection, coding, and multilevel signaling. Our results, for a simulated binary 8 Gbps system, show that simple techniques can be used to substantially reduce intersymbol interference, increasing system margin by several dB. A novel, but simple, nonlinear cancellation technique can more than *double* the dispersion-limited distance and/or data rate.

Nonparametric Inference for a Doubly Stochastic Poisson Process

Klaus Utikal *Department of Statistics, University of Kentucky, Lexington, KY 40506-0027*

Consider a doubly stochastic Poisson process whose intensity λ_t is given by $\lambda_t = \alpha(Z_t)$, where α is a unknown nonrandom function of an information process Z_t. Only one continuous time observation of counting and information process is available. The function $A(z) = \int_0^z \alpha(x)\,dx$ is estimated from the class of Lipschitz continuous functions α. The normalized estimator is shown to converge weakly to a Gaussian process as time approaches infinity. Confidence bands for A are given. Tests for independence from the process Z_t are proposed.

Analysis of Coherent Random-Carrier CDMA and Hybrid WDMA/CDMA Multiplexing for High-Capacity Optical Networks

B. Ghaffari and E. Geraniotis *Department of Electrical Engineering & Systems Research Center, University of Maryland, College Park, MD 20742*

In this paper we provide an exact analysis of the performance of a random-carrier (RC) code-division multiplexing (CDMA) scheme recently introduced for use in high-capacity optical networks. According to this scheme coherent optical techniques are employed to exploit the huge bandwidth of single-mode optical fibers and are coupled with spread-spectrum direct-sequence modulation in order to mitigate the interference from other signals due to the frequency overlap caused by the instability of the carrier frequency of the laser.

The average bit error probability of this multiplexing scheme is evaluated with arbitrary accuracy by integrating the characteristic functions of the components at the output of the matched optical filter

due to the desired signal and the other-user interference. Both phase noise and thermal noise (AWGN) are taken into account in the computation. The analysis is valid for any spreading gain and any number of interfering users and does not make use of limiting theorems and Gaussian approximations. The performance evaluation of RC CDMA establishes the potential advantage in employing hybrids of frequency-division and code-division multiplexing to combine the best features of both multiplexing schemes. (This research was supported in part by the Office of Naval Research under contract N00014-89-J-1375 and in part by the Systems Research Center at the University of Maryland, College Park, through the National Science Foundation's Engineering Research Centers Program: NSF CDR 88003012.)

Soft Maximum Likelihood Detection for Balanced Binary Block Codes

Michael Hall and Gareguin S. Markarian *Communications Laboratory, Helsinki University of Technology, Otakaari 5A, SF-02150 Espoo, Finland, and Radiophysik and Electronic Institute, Armenian Academy of Sciences, 378410, Ashtarak-2, Armenia, USSR*

In this paper, a simple and fast soft decision decoding algorithm is proposed. The algorithm reduces the probability of error per code word in balanced nB-(n+1)B codes, which are often utilized in digital fiber-optic transmission systems generally employing symbol-by-symbol hard decision, without greatly increasing the complexity of the receiver.

Using a 1B-2B and 3B-4B code as an example, it is shown that the algorithm is simple to implement in practice. The soft decision algorithm decoder of the 3B-4B code can be designed to be implemented using only relatively few comparisons, as compared to other soft decision decoding algorithms. The operation of the 3B-4B decision algorithm is presented in detail in the paper.

SESSION TA4

SOURCE CODING I

A New Example of Optimal Source Coding for Infinite Alphabets

Julia Abrahams *Mathematical Sciences Division, Office of Naval Research, Arlington, Virginia, 22217-5000*

The Huffman code is found for a particular integer source. The probability distribution of the source alphabet is given by

$$p_i = (1-\theta)(\theta/2)^{[\log_2 i]}, \qquad i=1,2,\cdots$$

for $0.390 \le \theta \le 0.618$. The optimal code has a length vector given by

$$l_i = 2[\log_2 i] + 1, \qquad i=1,2,\cdots$$

($[x]$ is the largest integer less than or equal to x.)

The method appears to be amenable to other parametric families of distributions. For a parametric family which includes a highly structured dyadic distribution, the length vector can be determined from that dyadic distribution. Numerical evidence suggests the form of the code tree to be verified as optimal by means of Gallager's sibling property. Ideally the sibling property is equivalent to a few inequalities that determine the applicable parameter range. This approach extends the class of integer alphabets for which optimal binary codes are known. Previously only the geometric and Poisson alphabet cases were solved.

Optimization of Overlapping Block Transform for Source Coding

Miodrag Temerinac and Bernd Edler *Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover (FRG)*

Transform coding with or without block overlapping (TC) and subband coding (SBC) are two methods for bit rate reduction of correlated signals. In this paper a common theory for both of these methods is applied, which is called "L into N coding" (LINC), since both methods map L input signal samples into N output coefficients. Several realizations of LINC with exact reconstruction properties are known. They are either restricted to certain values of N or L, or they need complex design and optimization strategies. For example Conjugate Quadrature Filters (CQF) and Generalized Quadrature Mirror Filters (CQMF) are restricted to $N = 2$ and higher values of N are only possible by cascading these structures. The transform with Time Domain Aliasing Cancellation (TDAC) and the Lapped Orthogonal Transform (LOT), on the other hand are restricted to $L = 2N$.

It is shown in this paper, that the overlapping block transform (OBT) is a suboptimal solution of LINC with arbitrary values of L and N. Furthermore, a fast algorithm and a simple optimization strategy based on effective bandwidth minimization are presented. The proposed optimization method for OBT can be done independent of signal statistics and provides a high coding gain. The optimization results for $2 \le N \le 32$ and $N \le L \le 8N$ are analyzed and compared with those of a DCT on one hand and theoretical bounds on the other hand. The OBT provides a coding gain, which is higher than that of DCT, is independent of the correlation sign, and almost reaches the theoretical bounds.

On Entropy Rate for Source Encoding on a Pyramid Structure

R. P. Rao and W. A. Pearlman *Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180*

Entropy is defined in terms of the source spectrum and is used to analyze pyramid structures. Two different pyramid structures are considered - the difference pyramid and the orthogonal pyramid.

Conditions, in terms of spectral entropy, are derived under which a pyramid structure is optimum. The Laplacian pyramid is also studied as a special case of the difference pyramid and theoretical justification for its optimality is given. The orthogonal pyramid has the property that spectral entropy of the pyramid is equal to that of the full-band signal and this property is used to prove that the two are equivalent in the rate-distortion sense. It is also shown that the combined first-order entropy of an optimally filtered pyramid is closer to the entropy rate of the source than the first-order entropy of the full-band signal. This means that loss-less compression of the source sequence can be obtained by merely representing it on a pyramid structure. Experimental results using both synthetic sources and speech samples are presented to verify the above claims.

A Combined Source and Error Correcting Code

D. Taipale *Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1084*

When signaling over some channels (such as on-off keying on a fiber optic link), the error probability can be reduced by reducing the a-priori probability of a pulse. In systems which use error correcting coding, this is difficult because most error correcting codes usually have as many ones as zeros (on the average). In this paper, we develop an algorithm which can be combined with any linear, cyclic code to obtain a combined source and error correcting code with any specified average symbol probability distribution. The main idea in the algorithm is simple, each codeword is used as a distinct symbol in source code output alphabet. The source encoding algorithm we use is a practical form of the Elias code developed by Jones. This algorithm requires that we order the codewords, and that for each codeword, we know the total number of ones in the codewords that occur earlier in the ordering (the weight). For codes with a large number of codewords, it is not convenient to determine the weight using table lookup. To overcome this problem, we develop a way to find the weight (and so compute the correct codeword) one coordinate at a time.

Universal Source Coding with Order Information

Kenneth Keeler *Harvard University, Pierce Hall, Cambridge, MA 02138*

Consider the class $\Omega(HA, \leq_\bullet)$ of iid sources (characterized as probability distributions p) of symbols from a countably infinite partially ordered alphabet $(HA, \leq_\bullet)$ which obey the order "$\leq_\bullet$" in the sense that $x \leq_\bullet y \Rightarrow p(x) \geq p(y)$. (Decreasing distributions on the positive integers $(N, \leq)$ are the most basic examples.) This paper addresses the problem of developing efficient source codes for A given only that the true source p is in $\Omega(HA, \leq_\bullet)$, generalizing the work of Elias and Rissanen on representations of and universal codes for $\Omega(N, \leq)$. A code is considered to be universal if its length function ℓ minimizes the worst-case inverse coding efficiency $E_p\ell / H(p)$ over $p \in \Omega(HA, \leq_\bullet)$ in the high-entropy limit. The existence, behavior and performance of universal codes for $(HA, \leq_\bullet)$ are expressed in terms of the asymptotic behavior of the predecessor cardinality $P(x) = |\{y : y \leq_\bullet x\}|$ and the cardinality of its level sets; the characterization is complete for the case of weak orders. These results are of particular importance in methods of inference by code length minimization, their application to which is discussed with illustrative examples.

Combined Source-Channel Coding for Band-limited Waveform Channels

V. Vaishampayan and N. Farvardin *Electrical Engineering Department, Systems Research Center, and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742*

In this paper, we consider the problem of optimum block-structured communication system design for a band-limited additive white Gaussian noise channel. The transmitter consists of a block encoder, a signal selection unit and a modulator. The receiver consists of a demodulator and a source decoder. The objective is to design the source encoder, the modulation signal set and the decoder so as to minimize the mean squared-error subject to constraints on the average transmitted power and the channel bandwidth.

Necessary conditions for optimality are derived; however, it is difficult to solve these necessary conditions in the most general case. For the special case where the modulation signal set belongs to the QAM family and for a maximum-likelihood (ML) demodulator, we have developed an algorithm that iteratively solves the necessary conditions for optimality for the encoder and the decoder subject to the above mentioned constraints. Numerical results are obtained for Gauss-Markov sources and different choices of system parameters. For comparison purposes, we have also determined the performance of a conventional vector quantization system combined with QAM modulation and ML demodulation. These results indicate that significant performance improvements over the conventional vector quantization system can be obtained. Furthermore, comparisons are made against a system with a linear decoder in which the encoder, signal set and the linear decoder are optimized. Finally, various channel mismatch issues are studied and the relative robustness of the above systems are investigated.

Source and Channel Entropy Coding

George H. Freeman *Department of Electrical Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

Entropy codes are classified into two kinds. Source entropy codes such as common string parsers alter the source statistics to match a given channel. Channel entropy codes such as the Huffman codes alter the channel statistics to match a given source. The two kinds of codes are given dual descriptions using parse/code trees. An heuristic and possibly optimal algorithm for the joint design of the trees yields codes for the binary i.i.d. source having good noiseless compression with complexity much lower than either parsing or Huffman coding alone. The number of possible tree configurations is limited by the number of Huffman code trees. Recursive expressions are derived for the number of Huffman code trees with given height or given total number of nodes.

Source Coder Structural Constraints and Information Patterns

Jerry D. Gibson, Thomas R. Fischer, and Wen-Whei Chang *Department of Electrical Engineering, Texas A&M University, College Station, TX 77843; Department of Electrical & Computer Engineering, Washington State University, Pullman, WA 99164; and Department of Communication Engineering, National Chiao-Tung University, Taiwan, R.O.C.*

Papers by Fine and Gibson and Fischer provide an approach for the design of data compression systems subject to a constraint on the transmitted alphabet. One step in this procedure employs what has been called the minimum search property. Gabor, *et al.* claim that this minimum search property may not be satisfied by optimal encoder/decoder pairs and present a purported counterexample with certain structural constraints. We show that the minimum search property is satisfied by encoder/decoder pairs with classical information patterns, as required in the theory presented by Fine and by Gibson and Fischer, and that the counterexample is invalid since it has a non-classical information pattern imposed by the structural constraints. We then solve the example with a classical information pattern and the minimum search property.

SESSION TA5

SHANNON THEORY II

Diversity Coding for Transparent Self-Healing and Fault-Tolerant Communication Networks

Ender Ayanoglu, R. D. Gitlin, Chih-Lin I, and J. E. Mazo *Communication Systems Research Laboratory, AT&T Bell Laboratories, Crawfords Corner Rd., Holmdel, NJ 07733-1988, and Mathematical Sciences Research Center, AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-2070*

In this paper, a channel coding approach called *diversity coding* is introduced for self-healing and fault-tolerance in digital communication networks for nearly instantaneous recovery from $n \leq M$ link failures. To achieve this goal, the problem of link failures is treated as an erasure channel problem. Two methods are presented: (i) the parity generator matrix of the code is Fourier, i.e., the parity lines are a discrete Fourier transform of the data in $GF(2^m)$, (ii) the parity check matrix of the code is equivalent to a Fourier matrix, i.e., the code is Reed-Solomon. For a given m, the first code always requires fewer additions and multiplications, and for $m = 2$, and 3, it also requires a smaller m, achieving the smallest field size possible. The simple, point-to-point technique is then extended to an arbitrary network topology. Implementation details of this technique in existing and future communication networks are discussed, and applications of the technique to trunk failures, short-term environmental disruptions such as fading channels in microwave radio networks or polarization dispersion in fiber optic networks with wavelength division multiplexing, dispersity routing in packet-switched networks, distributed storage, and fault-tolerant parallel transmission of continuous-amplitude discrete-time signals are presented.

What is the Capacity of One Bit of Memory?

Santosh S. Venkatesh *Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104*

We investigate the minmax problem of the maximization of the minimum amount of information that a single bit of memory retains about the entire past. Specifically, we are given a random binary sequence of $\pm$ inputs drawn from a sequence of symmetric Bernoulli trials, and a family of (time dependent, deterministic or probabilistic) memory update rules which at each epoch produce a new bit (-1 or 1) of memory depending solely on the epoch, the current input, and the current state of memory. The problem is to estimate the supremum over all possible sequences of update rules of the minimum information that the bit of memory at epoch n retains about the previous n inputs. In this paper we show precise estimates and demonstrate that a sequence of probabilistic memory updates--the *harmonic update rule*--retains the maximum amount of information about the past. We also investigate natural generalizations of the problem when more than one bit of memory is available.

Minimum Bound of Auto- and Cross-Correlation of Sequences

Shuo-Yen Robert Li, and Ning Zhang *Bellcore, Morristown, NJ 07960-1960 and Pacific Bell, 2600 Camino Ramon, 1S900G San Ramon, CA 94583*

In some communications systems, a receiver uses matched filter detection for a desired signal. The detector outputs the correlation function of the input signal with the desired signal. It is therefore preferable that the signal has an impulse-like autocorrelation function. When multiple signals are used in the same channel, it is also preferable that the crosscorrelation function between two different signals is small in magnitude. We investigate the minimum bound on all crosscorrelations and all off-phase autocorrellations of two signals, which are equally long sequences of unit vectors. In the special case when the two signals are generalized Barker sequences, we also investigate the minimum bound on all their crosscorrelations.

Binary Quadratic Form: a Solution to the Set Partitioning over $GF(q)$

Celso de Almeida and R. Palazzo, Jr. *Department of Telematica, FEE-UNICAMP, P.O. Box 6101, 13081 Campinas, SP, Brazil*

In this paper it is presented a general solution to the set partitioning problem over $GF(q)$ by use of the binary quadratic forms for bi-dimensional lattices. From Fermat's results. Genus and Composition Theorems, it is shown that when the solution is relatively prime the resultant fundamental set form a Latin Square and that the least squared Euclidean distance between points belonging to the same coset is q.

An Extended Cutoff Rate for Frequency-Hopping Communications with Non-Ideal Interleaving

Shaul Laufer and Jakov Snyders *Department of Electronic Systems, Tel-Aviv University, P.O. Box 39040, Tel-Aviv 69978, Israel*

The suitability of the capacity C and the cutoff rate R_o, which are the customary measures for evaluating the channel's ability to transmit information reliably, turns out to be questionable with respect to frequency-hopping communications over channels with partial band jamming and non-ideal interleaving. More explicitly, non-ideal pseudo-random block interleaving is assumed, resulting in a block interference channel with noise severity level which depends on the number of jammed hops in each block of interleaving. As the interleaving span increases, R_o for this channel departs, inadequately, from the cutoff rate for the memoryless channel. Also, C does not reflect the behavior of the channel when short block codes, those usually encountered in interleaved communications, are used.

An extended cutoff rate R_{oe} for channels with block memory is introduced. It is given by

$$R_{oe}(N,N_c) = \min_{\rho}\left\{-\frac{1}{N_c}\log_M E\left[M^{-N_c\cdot R_o(n)}\right]\right\}$$

where N, N_C, ρ, are the interleaving span, codeword length and fraction of band jammed, respectively, $R_o(n)$ is the cutoff rate for a pseudo-randomly interleaved block with n jammed hops and the expectation is taken with respect to n. Expressions for the extended cutoff rate are derived, both in presence and absence of side information. R_{oe} has distinctive properties. In particular, as the interleaving span tends to infinity, R_{oe} converges asymptotically and apparently monotonically to R_o for the memoryless case.

When Do Three Convex Corners Generate the Unit Simplex?

K. Marton *Mathematical Institute of the Hungarian Academy of Sciences, Budapest, P.O.B 127, H-1364 Hungary*

Let R_k^+ denote the non-negative orthant of the k-dimensional Euclidian space. A set $A \subset R_k^+$ is called convex corner if it is convex, compact, has non-empty interior, and $a \in A$, $a' \le a$ imply $a' \in A$. ("$a' \le a$" is understood coordinatewise.) The unit corner, or unit simplex, T is the simplex with vertices (0,...,0), (1,0,...,0), (0,1,0,...,0),...(0,0,...,1). The antiblocker of the convex corner A is the convex corner

$$A^* = \{b \in R_k^+ : \quad (a,b) \le 1, \quad \text{all} \quad a \in A\}.$$

For two vectors $x,y \in R_k^+$, $x \circ y$ denotes the vector $(x_i y_i, \quad i = 1,...,k)$. It has been shown recently that for two convex corners A and B, satisfying $B \subset A^*$, we have

$$T = A \circ B \text{ iff } B = A^*.$$

In the present paper we give, under some restrictions, a necessary and sufficient condition for three convex corners A, B and D, to satisfy

$$T = A \circ B \circ D.$$

An Algorithm for the Piecewise Linear Approximation of Planar Curves

Shuichi Itoh *University of Electro-Communications, Chofu, Tokyo 182, Japan*

The approximation of two dimensional digital curve is treated within a framework of universal data compression. The given curve is supposed to have random fluctuations. Our object is to get a smooth approximation of it by a piecewise linear curve. We assume that the given curve is generated by a noisy observation of the desired approximation. According to this source model, those points on the given curve are encoded by first describing the approximation and then by describing the deviation of those points. Following Rissanen's Minimum Description Length Principle, the approximation yielding the shortest description of those points is formulated. An algorithm is proposed to solve the above problem by a restricted optimization, whose computational complexity is of $O(N \log N)$ where N is the number of the points on the given curve. The algorithm is robust in a sense that no tolerance parameter is needed.

Informatic Crossover in Genetic Algorithms

Sami Khuri *Dept. of Biomedical Eng., The Johns Hopkins University, Baltimore, Maryland*

Genetic algorithms are randomized (but not directionless) search procedures that maneuver through complex spaces looking for optimal solutions. They are based on the mechanics of natural selection and natural genetics. These procedures, which are implicitly parallel in structure, make few assumptions about the problem domain and can thus be applied to a broad range of problems.

A succession of generations created by reproduction, crossover and mutation constitutes a parallel search through the search space, favoring regions with above average fitness. The crossover operation is of vital importance since a loss of diversity in a generation will generally yield premature convergence, pushing the search toward a sub-optimal solution. This paper presents "informatic crossover", which is based on Shannon's information theory. It uses conditional entropy for selecting the crossing site in crossover and has been shown to promote diversity in generations.

SESSION TA6

CODING THEORY III

General Soft Decoding of Block and Convolutional Codes

John T. Coffey and Rodney M. Goodman *California Institute of Technology, Pasadena, CA 91125*

We examine the complexity of general methods for decoding linear block codes with full hard decision decoding and (especially) bounded soft decision decoding. We compare the complexity of information set decoding with that of other algorithms, and show how improvements are possible in the case of bounded soft decision decoding. Finally, we discuss the applicability of these results to the decoding of convolutional codes.

On the Enumeration and Generation of Non Weight Equivalent Rate ½ Convolutional Codes

Jean Conan and Chahin Fiouzi *Département de Génie Electrique, Ecole Polytechnique de Montréal, P.O. Box 6079, station "A", Montréal, Quebec H3C 3A7*

In this paper we investigate the question of weight equivalence of rate ½ binary convolutional codes and consider the problem of enumeration and generation of all the non weight equivalent such codes for a given memory order *m*. In the generation process, this will be done by ordering the potential encoders into classes in such a way that it becomes easy to recognize the weight equivalence as well as the catastrophic error propagation conditions. Consequently, sub-classes including self and non self reciprocal as well as catastrophic and non catastrophic such encoders will be introduced. The cardinal number of these ensembles as well as the number of "non weight equivalent" codes will be computed recursively as a function of *m*. Since these relations amount to simple convolutions, they can be compactly represented by generating functions which will also be presented.

Convolutional Codes for Finite State Channels

Manju Hegde, Mort Naraghi-Pour, and Xiaowei Chen *Dept. of Electrical and Computer Engineering, Louisana State University, Baton Rouge, LA 70803*

Error control strategies for finite state channels often utilize decoders designed for use over memoryless channels ("memoryless decoders"). These are used with either specifically designed codes such as burst error correcting codes or with random error correcting codes in conjunction with interleaving. It may be possible, however, to improve the error probability performance of coding schemes by employing decoders which exploit the memory in the channel. Such decoders would utilize, directly or indirectly, the channel state sequence information contained in the channel output sequence.

The notion of state in convolutional codes and the existence of efficient decoding algorithms which are sequential (e.g., Viterbi algorithm) make convolutional codes suitable for adaptation for use with decoders which utilize the memory in the finite state channel without greatly increasing the concomitant complexity. In this paper we propose several new decoders for the use of convolutional codes over such channels. The main feature of these decoders is that they can be implemented sequentially on a trellis by a "Viterbi-like" algorithm. We evaluate the performance of these decoders by simulation and compare it to the performance of "memoryless decoders" with and without interleaving. Our results indicate that in cases where the channel is bursty, these decoders significantly outperform the "memoryless decoders". The cost of this improved performance is somewhat moderate increase in the decoders complexity.

A Method for Calculating Weight Distribution of $R = k/n$ Convolutional Codes

Hiroshi Sasano and Masao Kasahara *Faculty of Science and Technology, Kinki University, Higashi-osaka, Osaka 577, Japan, and Department of Electronics, Kyoto Institute of Technology,*

Convolutional codes are nowadays extensively used on the various communication systems. In order to estimate the performance of convolutional codes, weight distribution is one of the most important parameters. Unfortunately, for rate $R = k/n$ codes, it is not easy to obtain the weight distribution even with relatively short constraint length.

In this paper, an efficient method for calculating weight distribution of binary rate $R = k/n$ convolutional codes is presented. It extends the earlier work for $R = 1/n$ convolutional codes. The method is based on the idea of the variation of Hamming weight of codewords caused by input binary sequences of an encoder, and the rule of the variation defined by a variation and the input sequence which determines it. Weight of a codeword is expressed by the sum of the variations and the base-weight which is the weight of the codeword corresponding to the shortest input sequence. The weight distribution is obtained iteratively in order of weight of codewords. This method consists of several procedures. It is shown that the procedures are systematic and easy to be handled.

The Free Distance of Fixed Convolutional Rate 2/4 Codes Meets the Costello Bound

V. V. Chepoyzov, B. J. M. Smeets, and K. Sh. Zigangirov *Inst. for Problems of Information Transmission, USSR Academy of Sciences, Moscow, USSR; Dept. of Information Theory, University of Lund, P.O. Box 118, S-221 00 Lund, Sweden; and Inst. for Problems of Information Transmission, USSR Academy of Sciences, Moscow, USSR*

The free distance d_{free} of fixed convolutional codes of rate $R = 2/4$, is asymptotically lower bounded by the Costello bound on d_{free} for time varying convolutional codes of this rate. In particular, we can show that for these fixed codes we have $\frac{d_{free}}{\upsilon} = \rho_c - O(\frac{1}{\log(\upsilon)})$, where $\upsilon = 4(m+1)$ is the constraint length, m is the memory, and $\rho_c = -R/\log(2^{1-R} - 1) = 0.3932$ is the Costello parameter for rate $R = 2/4$ codes.

Design of Non-Systematic 3-SyEC/AUED Codes of Asymptotically Optimal Order

Sandip Kundu *IBM T.J. Watson Research Center, 24-258, P.O. Box 218, Yorktown Heights, NY 10598*

This paper considers the design of balanced binary block codes that are capable of correcting up to 3 symmetric errors and detecting all unidirectional errors. Non-systematic 3-Symmetric- Error-Correcting/All-Undirectional-Error-Detecting (3-SyEC/AUED) codes constructed in this paper are of asymptotically optimal order in redundancy. Little has been published on non-systematic t-Symmetric Error Correcting/All-Undirectional-Error-Detecting codes. 1-SyEC/AUED codes were constructed in 1977 and 1981, 2-SyEC/AUED codes in 1984 and 1988 [by the author].

Constructions and Bounds for Systematic and Nonsystematic tEC/AUED Codes

Frank J. H. Böinck and Henk C. A. van Tilborg *Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, The Netherlands*

Several methods to construct systematic and nonsystematic t-error correcting/all unidirectional error-detecting codes are described. The construction of nonsystematic tEC/AUED codes, presented here, makes use of existing constant weight codes or of block designs.

Systematic tEC/AUED codes can be made by adding a tail to a linear t-error correcting code, but other constructions are of an ad hoc nature. They will often be found as suitably chosen subsets of nonsystematic tEC/AUED codes.

Further bounds and extensive tables on the word length of systematic and nonsystematic tEC/AUED codes are presented.

A New Technique for Constructing $t-EC/d-ED/AUED$ Codes

R. Venkatesan and Sulaiman Al-Bassam *Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's Canada A1B 3X5 and Department of Computer Science, Oregon State University, Corvallis, OR 97331*

This paper presents a procedure to construct t-error correcting/ d-error detecting/all unidirectional error detecting ($t-EC/d-ED/AUED$) codes. This technique is similar to those recently reported to construct $t-EC/AUED$ codes. A $t-EC/d-ED$ code is chosen and then a tail is appended in such a way that the new code can detect more than d errors when they are unidirectional. The efficiency of the resulting $t-EC/d-ED/AUED$ codes compares favorably to the existing method. In addition, this procedure enables simple and fast encoding and decoding algorithms. An upper bound for the number of additional check bits required to incorporate unidirectional error detection capability is derived.

SESSION TA7

TRELLIS CODING II

Trellis Precoding

M. Vedat Eyuboğlu and G. David Forney, Jr *Codex Corporation, 20 Cabot Boulevard, Mansfield, MA 02048* (40 min.)

On a Gaussian channel with intersymbol interference and/or colored noise, trellis precoding is a method of combining the equalization performance of an ideal decision-feedback equalizer with the coding gain of known lattice-type coset codes, and in addition with substantial shaping gain. Trellis precoding is a generalization of trellis shaping to non-ideal channels, on the one hand, or of Tomlinson precoding to coded systems, on the other. In principle, on any linear Gaussian channel, trellis precoding can be used to approach the Shannon limit. The method is quite practical whenever channel information is available to the transmitter.

Erasure-Free Sequential Decoding and Its Application to Trellis Codes

Fu-Quan Wang and Daniel J. Costello, Jr. *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556*

The Erasure-Free Fano Algorithm (EFA), a Fano algorithm version of the Multiple Stack Algorithm (MSA) which is erasure-free in the sense that the buffer is guaranteed not to overflow if a large speed factor is used, is presented. Simulation results show that the EFA can perform as well as the MSA but requires a larger computational effort. Although the speed factor required to guarantee erasure-free decoding is large, the average computational load for both the EFA and the MSA is small. Then a modified version of the EFA, called the Buffer Looking Algorithm (BLA), is investigated. Simulation results show that the BLA needs a much smaller speed factor to guarantee erasure-free decoding than the EFA or the MSA while good performance is retained. The performance, complexity, and delay of the BLA are compared to the Viterbi algorithm. The BLA is shown to achieve equal or lower error probabilities with reduced complexity but a larger delay. The application of the BLA to the decoding of convolutional and trellis codes is discussed. (This work was supported by NASA grant NAG 5-557 and by NSF grant NCR 89-03429.)

Entropy-Constrained Trellis Coded Quantization

Thomas R. Fischer and Min Wang *Department of Electrical and Computer Engineering, Washington State University, Pullman, WA 99164, and Department of Electrical Engineering, Texas A&M University, College Station, TX 77843*

Entropy-constrained trellis coded quantization (ECTCQ) combines entropy coding with the trellis coded quantization (TCQ), yielding an encoding scheme that offers improved performance over entropy-constrained scalar quantization with only a modest increase in encoding complexity. For memoryless Gaussian and Laplacian sources, the ECTCQ system with 8-state trellis provides mean-square error (MSE) performance within about 0.6 dB of the respective distortion-rate function for encoding rates above about 1.5 bit/sample. Sources with memory can be encoded with a predictive ECTCQ system. For Gauss-Markov sources and encoding rates above 1.5 bit/sample, predictive ECTCQ with an 8-state trellis yields MSE within about 0.7 dB of the respective distortion-rate function. Good performance at lower encoding rates is achieved by using vector codebooks in the ECTCQ system.

The Design of Joint Source/Channel Trellis Coded Quantization/Modulation

Min Wang and Thomas R. Fischer *Department of Electrical Engineering, Texas A&M University, College Station, TX 77843 and Department of Electrical and Computer Engineering, Washington State University, Pullman, WA 99164*

A joint trellis coded quantization (TCQ) and trellis coded modulation (TCM) system is designed by selecting identical trellises in the TCQ and TCM, and mapping the quantization levels to modulation symbols. An algorithm is derived for the joint optimization of the TCQ levels and TCM symbols. Extensions of the approach allow predictive encoding and multidimensional codewords and modulation symbols.

Regular Labelings for Trellis Codes with Rectangular Signal Constellations

Ying Li *Dept. of EECS, MIT, Cambridge, MA 02139*

Regular trellis codes have the desirable property that the the minimum distance is relatively easy to calculate, and thus the code search and the performance evaluation are simplified. This paper presents new results on regular trellis codes using rectangular signal constellations. A "labeling" for a trellis code is a mapping from coded bits to subsets of signals in the partitioned signal constellation. A "regular labeling" is required for a trellis code to be regular. We study structures of regular labelings for rectangular signal constellations where the subset partitioning does not necessarily correspond to coset decomposition. It is shown that for an m-dimensional constellation partitioned into more than 2^{2m} subsets, regular labelings, and thus regular binary trellis codes, do not exist. For an m-dimensional 2^{2m}-way partition, it is shown that there is a unique structure for regular labelings. For m-dimensional partitions of fewer than 2^{2m} subsets, new structures for regular labelings are found.

Performance Bounds for Trellis Coded Direct Sequence Spread Spectrum Multiple Access Communications Systems

Brian D. Woerner and Wayne E. Stark *The Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109*

This paper considers a Direct Sequence Spread Spectrum (DS/SS) multiple access communications system. One may view a binary antipodal signature sequence as a low rate repetition code. Other choices for sets of signature sequences, corresponding to different codes, are possible.

We have previously proposed the use of trellis coded signature sequences as a means of improving the average probability of bit error for this type of communication system. The trellis code are obtained by performing set partitioning on a set of biorthogonal signature sequences. These codes have a coding gain over several decibels over the corresponding binary antipodal communications system.

We assume that the communications system employs a correlation receiver and a maximum likelihood (Viterbi) decoder. We apply a technique originally introduced by Lehnert and Pursley to analyze the performance of a DS/SS communication system with binary antipodal signature sequences. By slightly modifying this technique, we are able to obtain accurate performance results for our trellis coded systems. The results show that our trellis coded systems exhibit performance superior to that of binary antipodal DS/SS multiple access systems.

New Trellis Codes over $GF(Q)$ for One and Two Dimensional Lattices

Celso de Almeida and R. Palazzo, Jr. *Department of Telematics, FEE-UNICAMP, P.O. Box 6101, 13081 Campinas, SP, Brazil*

In this paper new trellis codes over GF(q) are presented for one and two dimensional Euclidean spaces. It is derived a closed form expression for the squared free Euclidean distance and for the asymptotic coding gain for the combined form of M-PAM with convolutional codes with one and two memory elements. The corresponding generator matrices for a class of these codes are also presented.

Curves showing the asymptotic behavior are provided. By use of the solutions of the diophantine equation (set partitioning) and the generalized minterm technique associated with the concept of Latin Square new trellis codes are tabulated for the combined form of QAM and Convolutional codes with one memory element.

TECHNICAL SESSIONS

Tuesday, 2 p.m. - 5 p.m.

SESSION TP1

STOCHASTIC PROCESSES

Extension of Slepian's Model of Gaussian Noise

Nelson M. Blachman *GTE Government Systems Corp., Mountain View, CA 94039-7188*

The notion underlying the Schmidt orthogonalization of polynomials can be applied to finite linear combinations of distinct variables. These variables may represent the value $x(t)$ and derivatives $x'(t)$, $x''(t)$, ..., $x^{(N)}(t)$ of a zero-mean stationary random process at a given instant. For this application the properties of the least-mean-square linear predictor of the random process are readily developed. The $N+1$ terms of this predictor are linear functions of $x(t)$ and increasing numbers of its derivatives; they are orthogonal to each other as well as to the error. By specializing to the Gaussian case, the Slepian model -- the sum of the first two terms plus its error -- is easily extended. The $N+1$ terms are deterministic functions of time that give the model the right conditional mean, and the $(N+2)$nd term (the corresponding error) is a zero-mean nonstationary Gaussian process that gives the model the proper conditional covariance. The details for a process with a Gaussian spectrum illustrate the results.

Sampling Designs for Estimating Integrals of Stochastic Processes

Karim Benhenni and Stamatis Cambanis *Department of Statistics, University of North Carolina, Chapel Hill, North Carolina*

The problem of estimating the integral of a stochastic process from observations at a finite number of sampling points is considered. Sacks and Ylvisaker (1966, 1968, 1970) found a sequence of asymptotically optimal sampling designs for general processes with exactly 0 to 1 quadratic mean (q.m.) derivatives using optimal-coefficient estimators, which depend on the process covariance. These results were extended to a restricted class of processes with exactly K q.m. derivatives for all $K = 0, 1, 2, \ldots$, by Eubank, Smith and Smith (1982). The asymptotic performance of these optimal-coefficient estimators is determined here for regular sequences of sampling designs and general processes with exactly K q.m. derivatives, $K \geq 0$. More significantly, simple nonparametric estimators based on an adjusted trapezoidal rule using regular sampling designs are introduced whose asymptotic performance is identical to that of the optimal-coefficient estimators for general processes with exactly K quadratic-mean derivatives for all $K = 0, 1, 2, \ldots$.

Almost Sure Convergence of Autoregressive Spectral Estimation

Elias Masry *Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093*

Let $\{X_n\}_{n=-\infty}^{\infty}$ be a stationary autoregressive process of order p with AR-parameters $\{a_j\}_{j=1}^{p}$, innovation variance σ_ε^2, and spectral density $\phi(\lambda)$. The observation process is $\{Y_n = X_n + W_n\}_{n=-\infty}^{\infty}$, where $\{W_n\}_{n=-\infty}^{\infty}$ is an additive white noise. On the basis of a finite length noisy observation $\{Y_n\}_{n=1}^{N}$, appropriate estimates of the parameters $\{a_j\}_{j=1}^{p}$, σ_ε^2, and the spectral density $\phi(\lambda)$ are considered and their almost sure convergence properties are derived. Sharp rates of almost sure convergence of these estimates are established (strong laws of the iterated logarithm). In particular, for the spectral density estimate $\hat{\phi}_N(\lambda)$, we have uniformly in λ, that

$$\left\{\frac{N}{(u_N \log\log N)}\right\}^{1/2} |\hat{\phi}_N(\lambda) - \phi(\lambda)| \to 0,$$

almost surely as $N \to \infty$, where $u_N \to \infty$ as $N \to \infty$ (at an arbitrarily slow rate).

Level-Crossing Analysis by Means of a Scaling-Dimensionality Transform

A. Barbé *Department of Electrical Engineering, Katholieke Universiteit Leuven, K. Mercierlaan 94, B-3030 Heverlee, Belgium*

It is well-known that the application of S.O. Rice's formula for the average number of level crossings over a certain time interval, gives the theoretically senseful but practically nonuseful result of this number being infinite in case the process considered is nondifferentiable. The problem in this contribution deals with a practicable characterization of the mean level-crossing activity of stationary processes.

With $E\,N_a(\theta)$ denoting the expected number of crossings of a fixed level a per unit of time, when the underlying process $x(t)$ is sampled and linearly interpolated with sampling period θ, the following function is defined: $D_a(r,\theta) = [\log E\,N_a(\theta) - \log E\,N_a(r\theta)]/\log r$. It is called the scaling-dimensionality transform and is somewhat related to the concept of fractal dimension. First, the properties of this transform are discussed in a general context. Then results will be given for stationary Gaussian processes for which the $E\,N_a(\theta)$ are easily calculable starting from Rice's formula. $D_a(r,\theta)$ is always finite, even for $\theta \to 0$, as opposed to $E\,N_a(\theta)$. For $\theta \to 0$, $D_a(r,\theta)$ corresponds to the fractal dimension of a related fractional Brownian motion as earlier considered by B. Mandelbrot. The concept of signal blurring is introduced, and its influence on the $D_a(r,\theta)$-transform is analyzed. This allows for a well-argued assessment of the level crossing activity of these signals, resulting in a well-balanced choice of both amplitude- and time-scale (sampling interval) for pixel-based signal representations.

The Curve Crossing Problem of a Gaussian Random Process

T. Munakata, T. Mimaki, and D.Wolf *Faculty of Engineering, Tamagawa University, Machida, Tokyo, Japan; University of Electro-Communications, Chofu, Tokyo, Japan; and Institut fuer Angewandte Physik, University of Frankfurt a.M., F.R.G.*

For a Gaussian random process having the autocorrelation function $m(\tau)$, the Rice function $Q^{\pm}(\cdot)$ of the curve crossing problems is reduced to a simple expression, if the curve has the form $C(\tau) = m(\tau)a + g(\tau)$, where a denotes the start level and $g(\tau)$ an arbitrary function. This expression $Q^{\pm}(\cdot)$ does not contain the value a explicitly. It means that the Rice function $Q^{\pm}(\cdot)$ is invariant for all curves defined above with given $\bar{g}(\tau)$ and arbitrary value of a. Furthermore, the following two special cases are very interesting:

CASE 1: for $g(\tau) = 0$, i.e., $C(\tau) = m(\tau)a$

The expression $Q^{\pm}(\cdot)$ becomes equal to the Rice function for the zero-crossing problem. This suggests that for the Gaussian process the property of curve-crossing intervals is invariant for any $C(\tau) = m(\tau)a$.

CASE 2: for $C(\tau) = m(\tau)a + (1 - m(\tau))b$.

$Q^{\pm}(\cdot)$ becomes equal to that for the level-crossing with level b. Now, one may question whether a similar discussion can be applied to the interval lengths between adjacent curve-crossing points. Since the curve crossing problem is not solved analytically, we made a computer simulation to answer the above question.

Multidimensional Random Fields Equivalent to Time Processes

Millu Rosenblatt-Roth *Department of Electrical Engineering, The City College of the City University of New York, New York, NY 10031*

The paper presents some main aspects of the theory of Markov random meshes: (a) The formal definition of the concept. (b) Because a Markov random mesh is a Markov field, it is necessary to express the characterizing elements of the later one as functions of the characterizing elements of the previous one. (c) Properties of the concept. (d) The homogeneous case. (e) The one-dimensional representation. (f) Examples of Markov meshes and of the corresponding Markov fields. (g) Examples of Markov fields which are not Markov meshes. The first example of such a random field appeared in 1965 (K. Abend, T.J. Harley, L.N. Kanal), but this idea was not pursued and quite forgotten. The author of this Abstract developed it to the level we can use it for modeling by best approximation of arbitrary random fields with the help of Markov meshes.

Constructing IVP Models with Specified Behavior on Certain Tail Events and Cylinder Sets

Amir Sadrolhefazi and Terrence Fine *School of Electrical Engineering, Cornell University, Ithaca, NY 14853*

We consider time-series models that can incorporate the seemingly contradictory properties of stationarity, continuity, and support of bounded yet divergent time averages. While the standard ergodic theorems rule out the existence of such models in the context of additive probability measures, there exist models with these properties in the framework of Interval-Valued-Probability (IVP). Formally, IVP is a pair of nonnegative and normalized set functions $(P,\bar{P})$, defined on a measurable space $(OMEGA, \mathsf{A})$, with P superadditive, $\bar{P}$ subadditive, and satisfying $P(\mathsf{A}) + \bar{P}(A^c) = 1$. We take our measurable space $(OMEGA, \mathsf{A})$ as $\Omega = \{0,1\}^N$, $\mathsf{A} = \sigma(\mathsf{C})$, where C is the field of all cylinders on Ω. Let D^* denote the divergence event (binary sequences with divergent averages), and define the dimension of a cylinder C, $dim(C)$, as the minimum number of coordinates required to specify C. We are interested in IVP models that can incorporate the following:

1) Stationarity
2) Continuity along C : $P(\cup_{i\geq 1} C_i) = \lim_n P(\cup_{i=1}^n C_i)$ and $P(\cap_{i\geq 1} C_i) = \lim_n P(\cap_{i=1}^n C_i)$ where $C_i \in \mathsf{C}$.
3) Support for divergence: $P(D^*) = \bar{P}(D^*) = 1$
4) Marginal constraints: Given an additive measure μ and a fixed integer k, $P(C) = \mu(C) = \bar{P}(C)$ for every cylinder C with $dim(C) \leq k$.

Specific constructions of IVP models that satisfy 1-3 have been given by Papamarcou and Grize; however, these constructions do not address the marginal constraints. We proceed to show that modulo the continuity condition, the issues of support for divergence and marginal constraints are unrelated for stationary IVP models. We provide sufficient conditions for the construction of IVP models that satisfy 1-4, and give specific instances where these sufficient conditions are satisfied.

Towards Robust Bispectrum Estimates

David J. Thomson *AT&T Bell Laboratories, Murray Hill, New Jersey 07974*

In many applications where the use of higher-order spectra is desirable, use of conventional methods is hampered by lack of data or, equivalently, by the evolutionary nature of the process. Estimates of bispectra are known to be highly variable and to exhibit anticonsistent behaviour. Here we study a robust variant of a multiple-window method for computing consistent estimates of the bispectrum from a short segment of the process. Bispectrum estimates are formed by robust time averages of triple products of what are effectively complex demodulates centered at f_1, f_2, and $-f_1 - f_2$.

SESSION TP2

MULTIPLE ACCESS I

Multiple-Access Coding with Error Control: A Code Construction for the Real Adder Channel

Peter Mathys *Department of ECE, Box 425, University of Colorado, Boulder, CO 80309*

A concatenated code construction for a noisy multiple-access channel (MAC) is given. The MAC which is considered is the binary input discrete-time real adder channel (DTRAC) with unknown gains, an unknown offset, and additive uncorrelated noise. It is assumed that the DTRAC is shared by a large number N of users of which at most T are active simultaneously. A nonlinear binary blockcode which is a subcode of a linear cyclic code (and thus easy to decode) over $GF(p_E)$, $p_E > 2$, is used for error control purposes. Linear binary codes of blocklength N are used for multiple-access coding. The latter codes have the property that the codewords of individual users which are active simultaneously can be separated in the generalized frequency domain over $GF(p_M^T)$.

Coding for the Multiple Access Channel with Sum-Rate Constraint

Bixio Rimoldi *Electrical Engineering Department, Washington University, St. Louis, MO 63130*

Consider M independent users sharing a discrete-time adder channel where the sum is taken over a finite field $GF(q)$. User i, $i = 1,...,M$, is allowed to transmit at the maximal rate R^i, but might choose to transmit at an effective rate R_E^i, $0 \le R_E^i \le R^i$. The problem is to find a linear block code C^i for each user i, such that the rate requirements are met and such that the transmitted information words can be recovered by the receivers provided that the (effective) sum-rate $R_s = \sum_{i=1}^{M} R_E^i$ satisfies $R_s \le R_{\max}$, where $R_{\max}$ is a fixed design parameter. Coding for both noiseless and noisy channels are discussed. A practical situation leading to the considered channel is encountered when M users want to transmit information across a single user channel with q-ary input and q-ary output: in this case an adder over $GF(q)$ can be inserted in front of the channel without decreasing the channel capacity.

Code Constructions for Asynchronous Random Multiple-Access to the Adder Channel

Eli Plotnik *Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

We consider a situation where up to T randomly chosen users, out of M potential ones, simultaneously transmit data over the noiseless Adder Channel. These T (or less) active users operate independently, and because of unknown time offsets among their clocks, and different delays that the various messages incur during transmission, both block and bit synchronism are precluded. Code constructions that ensure error-free asynchronous communication over this channel are presented. The information rate of these codes is $[T(1+\Delta/m)]^{-1}$ per user, where Δ is the maximal delay (measured in bits) between the transmission and reception times of a message in the system, and m is a free design parameter. If exactly T users are active, use of these codes leads to a stable throughput arbitrarily close to 1 message/slot. Furthermore, if the occurrence of collisions is made known to the active transmitters, such a throughput can be maintained for arbitrary T, $T \le M$ by means of an adaptive use of the codes.

Channel Capacity of Multiple-Access Channel with Binary Output

Yoichiro Watanabe *Stanford University, on leave from the Department of Electronics, Doshisha University, Kyoto, 602 Japan*

A sufficient condition for the channel capacity is proposed for a multiple-access channel with binary output. It is proved that a local maximum point of the mutual information gives the channel capacity. Since the mutual information of the multiple-access channel takes the value on the non-convex domain, the Kuhn-Tucker conditions concerning the mutual information is necessary conditions which determine the channel capacity. This is contrasted for the fact that the Kuhn-tucker conditions for the ordinary discrete memoryless channel are the necessary and sufficient conditions of determining the channel capacity. Thus the sufficient condition is desired to exactly determine the channel capacity. The channel capacity is indispensable to evaluation of the capacity region for a multiple-access channel.

The Capacity Region of the Random-Multiple Access Channel

Eli Plotnik *Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

The Random-Multiple Access Channel (R-MAC) is a communication system with M potential users, of which, T users or less are selected *at random* to transmit their messages simultaneously, over a common discrete memoryless channel. these T or fewer users are referred to as the *active* users, and they are fully synchronized. Based on the channel output vector, the decoder has to identify how many and which users are active, as well as their transmitted messages, with negligible probability of error. The capacity region of the R-MAC is defined as the set of all rate vectors $(R_1, R_2, ..., R_M) \in \mathbb{R}^M$ that ensure reliable communication.

Denote by X_l $(l = 1, 2, ..., M)$ the alphabet of user l and by Y the channel's output alphabet and let S_l $(l = 1, 2, ..., T)$ be the set of all subsets with exactly l users. When the users in the subset $s \in S_l$ are active, the channel is characterized by the transition probability measure $P(Y(\mid s)$. We prove the following theorem that establishes the capacity region of the R-MAC.

Theorem

Let $P(X_1, X_2, ..., X_M) \triangleq P_1(X_1) \cdot (P_2(X_2) \cdot \cdots \cdot P_M(X_M)$ be a probability distribution, and denote by $\Gamma(P_1, P_2, ..., P_M)$ the set of points $(R_1, R_2, ..., R_M) \in \mathbb{R}^M$ that satisfy the following conditions for *every* $I, J \in 1, 2, ..., M$ $(I \neq J)$:

$$R_I \leq \min \{ I(X_I;Y) \min_{l=1,2,...,T-1} \min_{\substack{s \in S_l \\ I \notin S}}$$

Differentially Coherent Multiuser Detection in Code-Division Multiple-Access Channels

Mahesh K. Varanasi *Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309*

The demodulation of differentially phase-shift keyed signals transmitted simultaneously via a CDMA channel is studied under the assumption of white Gaussian background noise. A class of bilinear detectors is defined with the objective of choosing the optimal bilinear detector. The optimality criterion considered is near-far resistance which denotes worst case bit error rate in low background noise over near-far environments. The optimal bilinear detector is therefore obtained by solving a minimax optimization problem and is shown to be a decorrelating detector. The bit error rate of the decorrelator for each user is equivalent to that of the corresponding optimum single-user DPSK detector in a reduced energy single user environment. A useful property of the decorrelator therefore is that the bit-error rate of each user is invariant to interfering signal energies and phases, thereby eliminating the near-far problem associated with the conventional single-user detection scheme. It is further shown that no other DPSK multiuser detector has a higher near-far resistance than does the decorrelator. That is, the optimally near-far resistant bilinear detector is optimally near-far resistant among all DPSK detectors.

Capacity of RMS Bandlimited Gaussian Multiple-Access Channels

Roger S. Cheng and Sergio Verdú *Department of Electrical Engineering, Princeton University, Princeton, NJ 08544*

Continuous-time additive white Gaussian noise channels with strictly time-limited and root mean square (RMS) bandlimited inputs are studied. RMS bandwidth is equal to the normalized second moment of the spectrum, which has proved to be a useful and analytically tractable measure of the bandwidth of strictly time-limited waveforms.

We show how the classical formulas for the capacity of strictly bandlimited single and two-user channels change under the RMS bandwidth measure. In addition, we consider channels where the inputs are further constrained to be Pulse Amplitude Modulated (PAM) waveforms. We analyze those channels under the assumption that the transmitters are symbol-synchronous and we find the pair of pulses that achieves the boundary of the capacity region. The shapes of the optimal pulses depend not only on the bandwidth but also on the respective signal-to-noise ratios.

The Effect of Coding on the Reliability of Computer Memories

Rajeev Krishnamoorthy and Chris Heegard *School of Electrical Engineering, Cornell University, Ithaca, NY 14853*

We consider the problem of the Mean Time to Failure (MTTF) of a random access memory which is protected by an on-chip single-error-correcting code. In a previous paper we studied the implementation of a code on a memory array and analyzed the effect on the yield (the probability that every bit on the memory array can be read correctly.) We restrict our attention to single-cell failures, since the purpose of coding is to combat single-cell defects.

We assume that cell failures form a Poisson process with parameter λ and that cell failures in a codeword form a Poisson process of intensity $n\lambda$. We take into account the event that not all codewords on the memory array are defect-free.

We derive an expression for the MTTF of a coded memory array and analyze its behavior as p varies, where p is the probability that a cell is defective at $t = 0$. As $p \to 0$ the MTTF of coded arrays is proportional to $N^{-1/2}$ (where N is the size of the memory), compared to N^{-1} for the uncoded case. Coding therefore provides two benefits: it significantly increases the lifetimes of the memories, and slows down the rate of decrease of the MTTF as N increases.

SESSION TP3

SIGNAL PROCESSING I

Predictive Contour Coding for an Object-Oriented Analysis-Synthesis Coder

M. Hötter and K. W. Hahn *Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover, Appelstrasse 9A, 3 Hannover 1, F.R.G.*

A predictive contour coding algorithm is presented which encodes the temporal differences of object shapes as generated in an object-oriented analysis-synthesis coder for moving images. Using stored shape information from the previous image, the object shape is approximated by a combination of polygon- and spline-representation. The quality of the contour approximation depends on the number of vertices which is adapted to the available transmission capacity, i.e., the smaller the data volume the coarser the approximation. By the presented algorithm, a high quality contour approximation can be achieved with an average bit rate of only about 0.3 bit per contour pixel. Therefore the presented predictive contour coding algorithm is appropriate for object-oriented coding of moving images at very low transmission rates where an efficient description and coding of the object shapes is needed.

The Extended Berlekamp-Massey Algorithm

Willard L. Eastman *The Mitre Corporation, Burlington Road, Bedford, MA 01730*

The Berlekamp-Massey algorithm solves a linear Hankel system of equations with a special right-hand-side. Since a Toeplitz system of equations can be reordered to convert it into a Hankel system, the algorithm can also be used to solve a Toeplitz system with a special right-hand-side. Toeplitz systems of equations arise in a number of important signal processing applications, but for some applications it is necessary to solve a system with an arbitrary right-hand-side. In this talk we extend the Berlekamp-Massey algorithm for synthesizing the shortest-length linear feedforward combinational circuit (LFCC) that generates a second given sequence t from the contents of the shortest-length linear feedback shift register (LFSR) generating a given sequence s. The LFCC synthesis algorithm can solve $p \times p$ linear Hankel or Toeplitz systems of equations with arbitrary right-hand-sides. Efficient VLSI implementations of algorithms for solving general Toeplitz systems are of great interest. Such an implementation is presented.

Complex Sequences over $GF(p^M)$ with a Two-Level Autocorrelation Function

M. Antweiler and L. Bömer *Institute for Communication Engineering, Technical University of Aachen, D-5100 Aachen, West Germany*

The paper proposes new complex sequences with elements which have constant absolute values of 1. The periodic autocorrelation functions of these sequences are shown to be two-level. Such sequences play an important role in synchronization, radar applications or in code division multiplex systems. The sequences are generated by three consecutive mapping processes: the first one maps the elements of $GF(p^M)$, (p prime; M integer) into the elements of $GF(p^J)$, (J integer and a divisor of M); the second one maps the elements of $GF(p^J)$ into the elements of $GF(p)$. Finally the elements of $GF(p)$, which are integers in the range $[0, p-1]$, are transformed into the pth roots of unity which build the complex sequence. The first and the second mapping steps are expressed in terms of the well known trace function. The linear span of the new sequences is examined and is proven to be larger than the linear span of complex m-sequences, if the parameters of the mapping processes are appropriately chosen. One realization of a sequence generator is proposed, consisting of shift registers, adders and a read-only memory (ROM).

Construction of a New Class of Higher-Dimensional Legendre- and Pseudonoise-Arrays

L. Bömer and M. Antweiler *Institute for Communication Engineering, Aachen University of Technology, Aachen, West Germany*

In this paper a new construction method for synthesizing two- and higher-dimensional Legendre arrays is introduced. The resulting arrays are q-dimensional arrays and every dimension has the size p, where p denotes an odd prime. The first element of the q-dimensional array is of value "0". All other elements are of value " -1" or " +1". With a theorem it is shown that all sidelobes of the periodic autocorrelation function of these arrays are " -1". If q is set to 1, the well known Legendre sequences are constructed.

If the first element of these arrays are replaced by " -1" or " +1", a new class of binary arrays is generated. It is shown that these arrays are pseudonoise arrays, if the number of elements N equals 3 mod 4. For N equals 1 mod 4 the periodic autocorrelation function sidelobes are either + -1" or " +3".

Multiple Bases Signal Representation, Coding, and Reconstruction

A. A. (Louis) Beex and Felix G. Safar *Bradley Department of Electrical Engineering, Virginia Polytechnic Institute & State University, Blacksburg, VA 24061*

The problem of efficient signal communication at low data rates involves, in general, the encoding of the source for maximum data compression at the transmitter end, and the reconstruction at the receiver end, from the received information and all the available a priori or side information. We propose an adaptive signal representation scheme, based on the use of multiple orthogonal basis sets, that exhibits very good potential in both the source encoding and the signal reconstruction end of the above problem. Our representation leads naturally to the splitting of the signal into additive components, each of which has a simpler description than the original process. In addition, it exhibits a structure similar to that of codebook based coding. As a result, a very compact signal representation can be achieved. A splitting procedure called recursive residual projection is proposed, and its performance evaluated for the separation of imagelike line signals into basis-defined "edge" and "texture" components. The coding of these components leads to lower rates than those for transform coding methods. In reconstruction, the representation can be considered as a well-behaved constraint. This allowed for the development of the corresponding unique projection operator, applicable for iterative reconstruction methods in general. In particular, we also propose a noise tolerant version of this operator, a so-called soft projection operator, capable of achieving convergence under noisy measurement conditions. Computer simulations of the representation, coding, and reconstruction aspects corroborate the usefulness of this proposed representation.

Fast Vector Quantization Algorithm by Using an Adaptive Search Technique

Kohji Motoishi and Takesi Misumi *Department of Computer Science and Communication Engineering, Kyushu University, Fukuoka 812, Japan*

This paper presents a fast vector-quantization algorithm that yields an optimal code at a small number of searches by using an adaptive search technique. In preparation, we made a table of distances between codes and decide the first probing code on the basis of a given codebook. When an input vector enters, the distance between the input vector and the first probing code is calculated. Comparing the calculated distance with the distance table, some codes are automatically eliminated from candidates for the optimal code. The efficiency of the algorithm, that is, the number of eliminated codes, is greatly influenced by how to select the probing code. The first probing codes is decided in advance on the basis of statistical information on the training sequence. The second and following probing codes are adaptively selected so that the number of eliminated codes becomes larger. The result of a preliminary experiment showed that the number of search was reduced to 1/11.7 in comparison with the full search algorithm and to 1/2.12 in comparison with the tree search algorithm.

An Analysis of Cepstrum Via Wave Moments

Anil Khare and Toshinori Yoshikawa *Department of EE System Engg., Nagaoka University of Technology, Kamitomioka-Machi, 1603-1, Nagaoka-Shi, Niigata-Ken 940-21, Japan*

Wave moments for a signal sequence and its corresponding cepstral sequence are related in a way which obviates the need for direct calculation of cepstral coefficients. Hence an ideal calculation for moment of cepstrum is possible even if the duration of cepstrum is infinite. Calculation of cepstral coefficients as such requires sophisticated phase unwrapping algorithms and is essentially prone to problems of aliasing. It is also possible to describe at least theoretically all the properties of a signal sequence from its wave moment. As an application of this, wave moments of minimum phase sequences are calculated from the corresponding linear phase sequences without actually calculating the minimum phase sequence.

On Model-Fitting for Fast-Sampled Data

Rajiv Vijayan and H. Vincent Poor *Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1101 W. Springfield Ave., Urbana, IL 61801*

The conventional discrete-time autoregressive model is poorly suited for modeling series obtained by rapidly sampling continuous-time processes. The extreme ill-conditioning of the covariance matrix to be inverted in such cases causes numerical instabilities in the Levinson algorithm for estimating the autoregressive parameters. An alternative model, based on an incremental difference operator rather than the conventional shift operator, has been developed recently by the authors jointly with Goodwin and Moore. As the sampling interval goes to zero, the parameters of this model converge to certain parameters which depend on the statistics of the continuous-time process. A Levinson-type algorithm can be employed for efficiently estimating the parameters of this model. In this paper, the properties of this and related difference-based algorithms are explored both analytically and numerically.

SESSION TP4

QUANTIZATION II

Asymptotics of Quantizers Revisited

László Györfi, Tamás Linder, and Edward C. van der Meulen *Hungarian Academy of Sciences, Technical University of Budapest, H-1111 Budapest, Stoczek u.2, Hungary, and Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, B-3030 Leuven, Belgium*

In this paper we reanalyze the results by Gish and Pierce (1968) on asymptotically efficient quantizing. In particular, asymptotic bounds are derived on the difference between the entropy of the uniform quantizer and that of the optimal quantizer when the mean square error becomes small. Hereby, no assumptions are made at all on the density of the random variable being quantized, and use is made of some classical results by Rényi (1959) and Csisár (1973). Also, following the work by Ziv (1985), non-asymptotic and distribution-free bounds on the difference between the entropy of the uniform quantizer and that of the optimal quantizer are derived if both have the same distortion. Finally, non-uniform quantizers are considered. For the latter case the asymptotic relation is investigated between the entropy of the quantizer and the entropy of the random variable being quantized, with no assumption at all on the density.

Low Dimension/Moderate Bit Rate Vector Quantizers for the Laplace Source

Peter F. Swaszek *Department of Electrical Engineering, University of Rhode Island, Kingston, RI 02881*

In this paper we present an *unrestricted vector quantizer* for the independent Laplace source $x(x \in \mathbb{R}^k)$, a source model useful for speech and image coding. The VQ is based upon Helmert's transformation

$$g = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} |x_i| \; ; \quad u_j = \frac{1}{\sqrt{j(j+1)}} \left[\sum_{i=1}^{j} |x_i| - j|x_{j+1}| \right], \quad j = 1, 2, \ldots, k-1$$

and consists of a scalar quantizer for g and a lattice-based uniform VQ on a dimension $k-1$ simplex, $\alpha_{k-1}(\hat{g})$, for the u_i. The VQ is unrestricted in that the resolution of the lattice VQ for the u_i varies with the result of the g quantization.

Asymptotic performance and design results are reviewed. Fischer's pyramid vector quantizer is shown to be a special case of our VQ; low-dimension results for the PVQ are provided. Implementation details of the VQ using various lattices are presented. Due to the geometric complexity of an exact analysis, an approximate design algorithm for finite bitrate is developed. Example designs with Monte Carlo performance simulations are presented for low dimension and moderate bitrates (2 through 4 bits per dimension) to demonstrate the utility of this approach.

Source/Channel Coding for Vector Quantizers by Index Assignment Permutations

Kenneth Zeger, Erdal Paksoy, and Allen Gersho *Dept. of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106*

An *index assignment function* for a vector quantizer in the presence of channel noise is a permutation of the quantizer codevector indices. Each output index from a vector quantizer is mapped by an index assignment function to a new index which is subsequently used as the input to a block channel coder. For a given vector quantizer (source coder) and a given error control block channel code, the choice of an index assignment function can have a profound effect on the average distortion achieved.

This paper shows that existing procedures for the design of index assignment functions in the zero redundancy case can be generalized to include the use of block channel coding of the permuted indices. An effective design method is introduced for constructing index assignment functions for arbitrary block

error control codes with the objective of minimizing average distortion. Extensive numerical results for first-order Gauss-Markov and speech sources and for systematic linear codes with varying code rates demonstrate substantial performance gain at various channel error rates. Often, a gap of several dB of signal-to-noise ratio can exist between a poor and a good index assignment function. One interesting consequence is that a good index assignment for a quantizer with no channel with no coder can often be superior in performance to a poor index assignment for the same quantizer using redundancy coding.

Optimal Quantization over a Finite-State Noisy Channel

Hong Shen Wang and Nader Moayeri *Wireless Information Networks Laboratory, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08855-0909*

The problem of quantizer design when the quantizer output is transmitted over a finite-state noisy channel is considered. Such a channel is a good model for fading channels, and the channel state information may be available to the receiver through signal-to-noise ratio measurements. Necessary optimality conditions are derived for the encoder when the decoder is fixed and given, and vice versa. In the absence of channel state information at the decoder, the channel is simply equivalent to a discrete memoryless channel, and previous results and methods are applicable. When side information is available, it can be used by the decoder along with channel outputs to better reproduce the source information. In either case, the optimality conditions lead to an iterative design algorithm. Numerical results are presented showing the performance of the system for various sources, channel models, and transmission rates. It is shown that side information may lead to noticeable improvements in performance.

Reduced Complexity Entropy-Pruned Tree-Structured Vector Quantization

Tom Lookabaugh, *Compression Labs, Inc., 1860 Junction Avenue, San Jose, CA 95134*

Vector quantization is a data compression technique in which a vector of source symbols is represented by a reproduction vector drawn from a finite codebook of such vectors. The index of the chosen vector is all that need be communicated. Since the probability distribution across codebook indices is not uniform, additional compression can be achieved by application of variable rate noiseless entropy coding to the indices.

Performance gains can be achieved if the vector quantizer and entropy code are designed jointly instead of sequentially. Entropy-pruned tree-structured vector quantization is a computationally attractive jointly designed system in which a large tree-structured vector quantizer is pruned so as to produce the smallest distortion among all pruned trees of the same or lesser entropy. In this paper, I show how binary arithmetic coding applied at each node of the tree search produces a system that is capable of progressive transmission and is also well suited to buffer-instrumented communication systems. In fact, the multiplication free Q-coder arithmetic code combined with a piecewise linear distortion function in the vector quantizer leads to a multiplication free encoder with good rate-distortion performance and significant flexibility. Simulations are performed for image coding experiments.

Optimal Scalar and Vector Predictive Quantizer Design

Amitava Ghosh and James George Dunham *Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275*

Differential Pulse Code Modulation (DPCM) systems are important systems for data compression of speech, images and other signals. An optimal DPCM system based on optimization theory and Lloyd-Max quantizers is developed. Its performance is compared to a matched DPCM system for first order Gauss-Markov sources with a mean square error performance criterion. The performance of the optimal DPCM system is found to be closer to the distortion rate bound at high bit rates than the conventional

DPCM. The theory is extended to the design of infinite-state vector predictive quantizers. Unlike the conventional memoryless and finite-state vector quantizers, whose design is based on long training sequence, a probability density function is used when designing infinite-state vector predictive quantizers.

The Asymptotic Distribution of the Error in Scalar and Vector Quantizers

Don H. Lee and David L. Neuhoff *Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109*

An approximate formula is derived for the probability density of the quantization error caused by scalar and vector quantizers with many quantization levels. The rth moment of this density yields Bennett's integral and its generalizations. Like Bennett's integral, the formula becomes asymptotically exact as the number of levels increases. An application to two-stage quantization is discussed.

Vector Quantizers Using Permutation Codes as Codewords

Luzheng Lu, G. Cohen, and Ph. Godlewski *Dept. of Mathematics, University of Toulouse, Le Mirail, 5, allées Antonio-Machado, 31058 Toulouse Cedex, France, and ENST, 46, Rue Barrault, 75013 Paris, France*

A new vector quantizer structure is proposed for reducing the coding complexity, memory requirements and the design computational burden. the Codebook of such a quantizer consists of the first words of M ($M \leq 64$) permutation codes. The encoding operation is then done by the minimum distortion word search and the fast permutation coding. the quantizer is uniquely defined by jointly specifying the first word set and the corresponding partition of the input space, so it can be constructed from an initial fixed first word set by a LBG Algorithm and by the simple permutation code design techniques. the simulation results for such new quantizers designed and tested on different signals are given.

SESSION TP5

CYCLIC CODES

New Bounds on Cyclic Codes from Algebraic Curves

J. Wolfmann *G.E.C.T., Université de Toulon, 83130, La Garde, France* (40 min.)

Starting from a deep link between the words of cyclic codes and plane algebraic curves over finite fields we use bounds on the number of rational points of these curves to obtain general bounds for the weights of cyclic codes.

Algebraic Decoding Beyond e_{BCH} of Some Binary Cyclic Codes

Jeanette Janssen *Facultad de Ingeniería, Prolongación Támpico s/n, Adpo. Postal 215, Suc "A", Salamanca, Gto 36730, Mexico*

It is well known that many cyclic codes have a true error-correcting capability e that is strictly greater than the error-correcting capability e_{BCH}, that follows from the BCH-bound. The well-known Berlekamp-Massey decoding algorithm decodes only up to e_{BCH} errors. A generalization of the Berlekamp-Massey algorithm exists that decodes in certain cases up to e_{Roos}, a lower bound on e coming from the Roos bound. M. Elia described an algebraic decoding algorithm for the binary Golay code that uses the full error-correcting capacity of this code. In this presentation similar, but sometimes more complex, algebraic decoding algorithms, that decode up to e' errors, $e_{BCH} < e' \leq e$, are presented for a number of other binary cyclic codes. One of these algorithms is derived in detail, to show the way in which these kinds of decoding schemes can be found.

Decoding Binary 2D Cyclic Codes by the 2D Berlekamp-Massey Algorithm

Shojiro Sakata *Department of Production Systems Engineering, Toyohashi University of Technology, Tempaku, Toyohashi 440, Japan*

Binary 2D cyclic codes are a two-dimensional generalization of binary cyclic codes. While binary cyclic codes are constructed in terms of the univariate polynomial ring $K[x]$ over the binary field $K := GF(2)$, binary 2D cyclic codes are in terms of the bivariate polynomial ring $K[x,y]$. In this paper, we present a method of decoding binary 2D cyclic codes by using the 2D Berlekamp-Massey algorithm which has been introduced as an extension of the Berlekamp-Massey algorithm to two dimensions, and discuss the error correcting performance of some 2D cyclic codes. As is the case with Stevens' approach of extending the BCH decoding procedure to cyclic codes more general than BCH codes, we need some trial and error, i.e., testing a certain number μ of field elements of an extension field $K := GF(2^M)$. We verify some merit of our approach by showing several simple examples of 2D cyclic codes. Some are not equivalent to any (1D) cyclic codes, and the other, which are equivalent to (1D) cyclic codes, have smaller values of μ than when they are decoded by the 1D Stevens method.

On Error and Erasure Decoding of Cyclic Codes

H. Shahri and K. K. Tzeng *Department of Electrical Engineering, Lafayette College, Easton, PA 18042, and Department of Computer Science and Electrical Engineering, Lehigh University, Bethlehem, PA 18015*

This paper presents procedures for error and erasure decoding of cyclic codes up to the Hartmann-Tzeng bound as well as to special cases of the Roos bound. The main part is on converting the problem of error and erasure decoding to an error-only decoding problem so that the Feng-Tzeng multisequence shift-register synthesis algorithms can be applied. This work has thus extended the result obtained by Stevens on error and erasure decoding of cyclic codes up to a special case of the HT bound.

Decoding of Cyclic Codes Beyond Minimum Distance Bounds Using Nonrecurrent Syndrome Dependence Relations

G. L. Feng and K. K. Tzeng *Department of Computer Science and Electrical Engineering, Lehigh University, Bethlehem, PA 18015*

The decoding capabilities of algebraic algorithms such as the Berlekamp-Massey algorithm, the Euclidean algorithm and out generalizations of these algorithms are basically constrained by the minimum distance bounds of the codes. Thus, when the actual minimum distance of the codes is greater than that given by the bounds, these algorithms cannot fully utilize the error-correcting capability of the codes. The limitation is seen to be rooted in the original Peterson decoding procedure followed by these algorithms which requi..s the determination of the error-locator polynomial from the Newton's identities which in turn require that the syndromes be contiguous resulting in a set or sets of linear recurrences. In this paper, we introduce a procedure which breaks away from this restriction and can determine the error locations from nonrecurrent syndrome dependence relations. This procedure employs a fundamental iterative algorithm, which we have introduced to derive the Berlekamp-Massey algorithm and its generalization, and an error-evaluation formula which is a generalization of Forney's. It can decode many cyclic codes up to their actual minimum distance an is seen to be a generalization of Peterson's procedure.

A Transform Based Decoding Algorithm for Cyclic Codes Via Non Preserving Permutations

R. M. Campello de Souza *Dept. of Electronics and Systems Communications Research, Group - CODEC, Federal University of Pernambuco, Cidade Universitária, 50741, Recife-PE, Brasil*

A new transform domain based decoding algorithm for cyclic codes is introduced. The technique is based on shortened syndrome look-up tables and makes use of preserving as well as of non-preserving permutations. The results presented in the paper are an extensior of some previous work done on the subject and although only binary codes are considered, the method can be generalized in a straightforward manner to nonbinary codes.

Metacyclic Codes

Roberta Evans Sabin *Computer Science Department, Loyola College, Baltimore, Maryland, and Computer Science Department, University of Maryland, Baltimore County, Catonsville, Maryland*

Since many of the most familiar linear error-correcting codes are ideals in group algebras, we wish to examine the structure of codes in group algebras based on non-abelian metacyclic groups. A code in such a binary group ring is either a two-sided or a one-sided ideal. When the group is of odd order, the ring is semi-simple and decomposes into a direct sum of minimal codes. Unlike the abelian case, however, the decomposition is not unique. Minimal components are isomorphic but are not necessarily combinatorially equivalent as codes. Like abelian codes, these codes have a simple algebraic structure which will aid in implementation. In most cases, these codes are found to be a significant improvement over abelian codes of the same length and dimension.

SESSION TP6

CODING THEORY IV

Linear Inequalities for Covering Codes

Zhen Zhang *Communication Sciences Institute, Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-0272*

This paper deals with $K(n,R)$, the minimal number of codewords in any binary code of length n and with covering radius R. This quantity has been previously studied by Cohen *et al.* and van Wee. We improve the lower bounds on $K(n,R)$ for over 100 pairs of n and R within the range $n \leq 33$ and $R \leq 10$ by proving linear inequalities satisfied by covering codes of the following form

$$l_C(\sum_{i=0}^{m} \lambda_i B_i) \geq \beta \mathbf{I},$$

where l_C is the indicator of the code C, $\mathbf{I}$ is the 2^n-vector $(1, 1, ..., 1)$ and B_i is a $2^n \times 2^n$ 0-1 matrix defined as follows

$$B_i = (b_{x,y})_{x \in \mathbf{F}_2^n, y \in \mathbf{F}_2^n},$$

where $b_{x,y} = 1$ iff $d(x,y) = i$ and 0 otherwise. Some improvements for the lower bounds on $t[n,k]$ are also obtained as by-products.

Covering Radius Problems and Character Sums

A. Tietäväinen *Department of Mathematics, University of Turku, SF-20500 Turku, Finland*

Using the Carlitz-Uchiyama bound for character sums Anderson found a lower bound for the minimum distance of the dual of a BCH code. In this paper an upper bound is given for the covering radius of that code and in fact of any code with large dual distance.

Lower Bounds for Binary Covering Codes

Iiro Honkala *Department of Mathematics, University of Turku, SF-20500 Turku 50, Finland*

We give some modifications of the van Wee lower bounds on $K(n,R)$, the minimum cardinality of a binary code of length n and covering radius R. We use results about the classical combinatorial problem of covering pairs by k-tuples. Our results give many improvements to the best previously known lower bounds on $K(n,R)$. We also study s-subjectivity and covering radius, and normal and subnormal codes.

On the Covering Radii of Codes over $GF(q)$

H. Janwa *Department of Mathematics, Caltech, Pasadena, CA 91125*

Let R denote the covering radius of a q-ary $[n,k,d]$ code. We give an upper bound on R (which resembles the Griesmer on n), with many examples of codes for which this bound is exact. For a class of codes, e.g., the optimal codes, we give an improvement of this bound. We give necessary and sufficient conditions for optimal codes to attain this bound. We present a lower bound on the algebraic geometric codes in terms of the rational points and the genus of the underlying curve and show that for many of these codes the lower bound coincides with our upper bound. These codes furnish many non-trivial examples of q-ary normal codes. We also give a criterion for a code to be abnormal. As an application, we give further examples of non-binary abnormal codes. If the code is cyclic with the generator polynomial $g(x)$, then we give an upper bound on R involving $g(x)$. Furthermore if $g(x)$ is irreducible, then upper bounds on R are given using Waring's problem in finite fields. We find exact values of

covering radius and minimum distance of some codes of this type including some 1-error-correcting quasi-perfect codes.

We give a generalization of the so-called "Norse bounds" for binary codes to q-ary codes. We conclude with a discussion of further work and a list of open problems.

Some Results on the Covering Radius of Codes

Xiang-dong Hou *Department of Mathematics, University of Illinois at Chicago, Chicago, IL 60680*

Let $t[n,k]$ be the smallest covering radius of any $[n,k]$ binary code. We determine the values or improve the lower bounds of several entries in the table of $t[n,k]$ ($n \geq 64$) by Graham and Sloane

$$t[19,9] = 4,\ t[38,6] \geq 13,\ t[39,15] \geq 8,\ t[40,6] \geq 14,\ t[42,15] \geq 9,$$

$$t[43,6] \geq 15,\ t[43,26] = 5,\ t[45,6] \geq 16,+t[47,7] \geq 16,\ t[51,33] = 5,$$

$$t[52,7] \geq 18,\ t[52,18] \geq 11,\ t[52,34] = 5,\ t[57,7] \geq 20,\ t[58,20] \geq 12,$$

$$t[59,8] \geq 20,\ t[59,44] = 4,\ t[62,36] = 7,\ t[64,8] \geq 22.$$

These are done through two inequalities which improve a result of van Wee.

We also show that for any binary linear code C with covering radius ≥ 3, its norm $N(C)$ introduced by Graham and Sloane satisfies

$$N(C) \leq 3R - 2.$$

This improves the upper bound of Adams by 3, and implies that C is normal if $R = 3$. (Normality means $N(C) \leq 2R + 1$.)

Joint Decoding and Phase Estimation Via the Expectation-Maximization Algorithm

Ghassan Kawas Kaleh *Ecole Nationale Supérieure des Telécommunications, Département Communications, 46, rue Barrault, 75634 Paris Cedex 13, France*

The Trellis-Coded Modulations are very sensitive to carrier phase offset. To get a significant part of the predicted coding gain, a reliable carrier reference is required. We present an iterative method for joint phase estimation and symbol decoding via the Expectation- Maximization (EM) algorithm. Carrier phase offset and noise variance estimation are based on the maximum likelihood criterion. Estimates are obtained via an iterative maximization of the Kullback-Leibler information measure. The Markovian properties of the encoder states sequences are used to calculate the required likelihoods. At the end of iterations, the likelihood functions calculated by the EM algorithm can easily give optimum decisions on information symbols which minimize the symbol error probability. The method can be applied to all codes that can be represented by a trellis. Simulation results will be presented.

Generalized Identity-Guards Algorithm for Minimum Distance Decoding of Group Codes in Metric Spaces

L. B. Levitin, M. Naidjate, and C. R. P. Hartmann *Boston University, College of Engineering, 44 Cummington Street, Boston, MA 02215*

A generalization of the zero-guard algorithm for any group codes in metric spaces with a group invariant metric is suggested. The algorithm makes use of a special subset G_0 of codewords, called identity-guards. Only these codewords (which can be precomputed) should be stored and used in the decoding procedure. In general, the number of identity-guards is a small fraction of the number of all codewords. Necessary and sufficient conditions for a set of codewords to be G_0 are found for the general case and, more specifically, for special cases, e.g. Strictly discrete spaces, soft decision decoding, arithmetic *AN* codes, q-ary linear codes, permutation codes.

On the Linear *M*-Algorithm

Harro Osthoff, Rolf Johannesson, Ben Smeets, and Han Vinck *Department of Information Theory, University of Lund, Box 118, S-221 00 Lund, Sweden, and Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

The linear complexity version of the *M*-algorithm (called *LM*-algorithm) which we presented at the ISIT88 in Kobe is investigated further. For binary rate $R = 1/2$ convolutional codes our algorithm extends in each decoding step M possible information sequence estimates. The M most likely of $2M$ estimates are found by searching for the median. The storage size needed to decode an information frame of length l is $2Ml$ bits and it is organized in such a way that our final estimate can be found by a simple trace back method. The main problem with the *LM*-algorithm is the correct path loss.

We have investigated different methods to recover the lost correct path. Several quick-look-in codes, non-systematic codes in feedback realization, and systematic codes with good profile spectrum were tested. These systematic codes show superior spontaneous path recovery.

We will also report on simulations of the *LM*-algorithm when it is used together with a systematic code for communication over an 8-level quantized AWGN channel. When we suspect a correct path loss the decoder is reset with the M most likely states.

SESSION TP7

ERROR-CORRECTING CODES I

Bounds on the Undetected Error Probabilities of Linear Codes for Both Error Correction and Detection

Mao-Chao Lin *Dept. of Electrical Engineering, National Taiwan University, Taipei 10764, ROC*

In this paper, we study the $(n, k, d \geq 3)$ binary codes, which are used for correcting every single error and detecting other error patterns over the binary symmetric channel. We show there exists one code such that its probability of undetected errors is upper bounded by $n+1)[2^{(n-k)}-n]^{-1}$. We also study codes of length n for correcting all the error patterns of weight at most λn and detecting other error patterns. We show that there exists an $(n, Rn, d \geq 2\lambda n + 1)$ binary linear code whose probability of undetected errors is upper bounded by $2^{-[1-R-H(\lambda)]n}$ as n approaches infinity, where $H(\cdot)$ is the entropy function and $1 - R > H(2\lambda)$.

A New Class of Random Error Correcting Codes

Sandip Kundu *Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598*

This paper considers the design of binary block codes that are capable of correcting up to 2 symmetric errors. These codes are based on similar equations as BCH codes, but their efficiencies differ. Efficiencies vary from matching the best known codes to being inferior to BCH codes. The construction principle is novel and the author contends that decoding is simpler.

Asymmetric Error Correcting Codes

Bella Bose *Department of Computer Science, Oregon State University, Corvallis, OR 97331-3902*

Non-linear but cyclic codes capable of correcting asymmetric errors are described. For these codes the syndromes directly give the symmetric functions of the error locations and so these codes are much easier to decode. The hardware implementation of the decoding algorithm is given. In many cases the information rate of these codes is as good as or better than the corresponding BCH codes.

On Codes Correcting/Detecting Symmetric, Unidirectional, and/or Asymmetric Errors

J. H. Weber, C. de Vroedt, and D. E. Boekee *Delft University of Technology, Department of Electrical Engineering, and Department of Mathematics and Informatics, 2600 Delft, The Netherlands*

A code is called t_1-SyEC t_2-UEC t_3-AsEC d_1-SyED d_2-UED d_3-AsED (with $t_1 \leq t_2 \leq t_3$, $d_1 \leq d_2 \leq d_3$, $0 \leq t_i \leq d_i$) if it can correct up to t_1 symmetric errors, up to t_2 unidirectional errors, and up to t_3 asymmetric errors, as well as detect from $t_1 + 1$ up to d_1 symmetric errors that are not of the unidirectional type, from $t_2 + 1$ up to d_2 unidirectional errors that are not of the asymmetric type, and from $t_3 + 1$ up to d_3 asymmetric errors.

In this paper we derive necessary and sufficient conditions for a code to be t_1-SyEC t_2-UEC t_3-AsEC d_1-SyED d_2-UED d_3-AsED. Hence we can obtain necessary and sufficient conditions for correction and/or detection of any combination of symmetric/unidirectional/asymmetric errors by making appropriate choices for t_i and d_i. This includes existing as well as new results that appear as special cases of this general result. For example, the conditions on a code to be t-SyEC d-UED (with $0 \leq t \leq d$) can be obtained by substituting $t_1 = t_2 = t_3 = d_1 = t$ and $d_2 = d_3 = d$. Further, it follows from the general necessary and sufficient conditions that some codes have stronger error correcting/detecting capabilities than they were originally designed for.

Theory and Construction of M-ary Error Correcting and Discriminating Codes

Kohichi Sakaniwa, Tae Nam Ahn, and T. R. N. Rao *Department of Electrical & Electronic Engineering, Tokyo Institute of Technology, Tokyo, and The Center for Advanced Computer Studies, University of Southwestern Louisiana, Lafayette, LA 70504-4330*

Hamming and Lee distance have been very well established and utilized to define error sets and to construct error control codes. However, as Berlekamp has correctly pointed out (in his book *Algebraic Coding Theory*), neither of these metrics fit the multi-level communication system using amplitude modulation under Additive Gaussian Noise (AGN) assumption.

Therefore, we first derive some generalized metrics (called *quasi-metrics*) defined for the n-tuples over the integer ring that are applicable to many actual channel models including multi-level amplitude modulation schemes with AGN. We then develop a new general class of error control codes, namely *error correcting and discriminating* codes. The error correcting and/or detecting codes become special cases of this general class of codes. In this context we establish another new concept called *error difference set*, E^*, and relate its maximum quasi weight $Q_{\max}(E^*)$ to the error correcting and discriminating capabilities of codes.

Finally, we give m-ary code construction techniques and establish the error control capabilities using the theory developed. In a second construction method, we show how the error discriminating property can be used to construct unidirectional error control codes using asymmetric codes.

A Construction Method for Multilevel Error-Correcting Codes Based on Absolute Summation Weight

Hajime Jinushi and Kohichi Sakaniwa *Department of Electrical and Electronic Engineering, Tokyo Institute of Technology*

By introducing the *generalized Hadamard matrix* (GHM), we present a new construction method for multilevel error correcting codes based on *absolute summation weight*. the proposed code has two special features: 1. Decoding is very simple, i.e., errors can be removed by first multiplying the received word by the inverse GHM and then rounding off the every digit of the transformed word. 2. Since we can get a variety of GHM's, multilevel error-correcting codes with various t (error-correcting capability) are easily obtained. Moreover, we show by computer simulation that a much better performance can be obtained by the transmission system employing the proposed code compared to the uncoded binary one.

On the General Error-Correcting Capability of Linear Codes

Hans-Andrea Loeliger *Institute for Signal and Information Processing, ETH Zentrum, CH-8092 Zürich, Switzerland*

Linear (n,k) codes over GF(q) are considered for use in correcting the error patterns e ($e \in \mathrm{GF}(q)^n$) in a specified set E of additive error patterns. A sufficient condition for the existence of a linear code that corrects all error patterns in E is derived; this condition depends only on the cardinality $|E|$ of E and generalizes the usual Gilbert-Varshamov bound for linear codes. By a simple union bound, it is further shown that, for arbitrary E and arbitrary probability distribution on the full set of error patterns, the average block error probability, conditioned on the event that the error pattern is in E, over all linear (n,k) codes satisfies $P_{B|E} < \frac{1}{2}|E|q^{k-n}$. By letting E be a set of typical error patterns, this result can be used to show that linear codes can achieve capacity on a broad class of additive channels.

Cluster-Error-Correcting Array Codes

P. G. Farrell *Electrical Engineering Department, University of Manchester, M13 9PL, UK*

Array error control codes are linear block and convolutional codes which are constructed from several single parity check or other component codes, assembled in two or more dimensions or directions, with emphasis on simple component codes and low complexity methods of decoding. Array codes are

particularly suitable for detecting and correcting two-dimensional bursts or cluster of errors. Clusters, patches or two-dimensional burst of errors occur in many digital transmission, processing or storage systems, wherever data is formatted in two dimensions (e.g., magnetic tape, etc.). Efficient array codes can be devised for correcting cluster errors in such two-dimensional data structures.

PLENARY SESSION

Wednesday, 8 - 8:50 a.m.

Routing in Interconnection Networks

Bruce Hajek *Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801*

Packet-switched interconnection networks, such as the one connecting 65,000 processors in the Connection Machine, offer the possibility of *programmable* connections between processors. Processors that are not directly connected communicate via multiple-hop routes, so that the processors appear to be connected at a higher logical level. Special purpose hardware design and a variety of dynamic control devices are used to gain speed. A similar style of packet switching has also been proposed for wide area networks. Traffic congestion must be made to quickly diffuse through the use of fast, local mechanisms. We will discuss mathematical and information-theoretic tools for analysis and modeling, with the goal of both understanding the general design problems and of suggesting new architectures for its solution.

We will focus on deflection routing, originally termed "hot-potato" routing, which is a technique for maintaining bounded buffers in a packet-switched communication network. If, due to congestion at a switch, not all packets can be sent out along shortest paths to their destinations, some packets are sent out on other links. The penalty is an increase in the distance traveled by packets, and the reward is the simplicity of switch design resulting from the absence of large buffers and routing tables. Traditional store-and-forward networks use extensive computation at the nodes to determine packet routes in order to use transmission bandwidth sparingly. In contrast, deflection leads to simple switches by making liberal use of transmission bandwidth.

Both the worst case and average case performance of deflection routing will be discussed. Approximate analysis of the transient and steady state behavior of deflection routing in hypercube and shuffle-like networks leads to a model in which the progress of a typical packet is described by a random walk on the network graph. The condition that packets reach their destination in spite of deflections is analogous to the condition of low-error probability for communication over a noisy channel. Bounds based on information-theoretic considerations are given on the average delay of a packet.

TECHNICAL SESSIONS

Wednesday, 9 a.m. - 12 m.

SESSION WA1

COMMUNICATION THEORY II

Non-Linear Sequences with Controllable Correlation and Complexity Properties

K. M. Ibrahim *Department of Electrical Engineering, University of Baghdad, Iraq*

The construction of non-linear pseudorandom (PN) sequences with controllable complexity and correlation properties is presented in this paper. Such sequences are widely used for communication systems to get higher data privacy and immunity against interference. The correlation properties of these sequences provide a simple method for synchronization between the transmitter and the receiver. The synchronization time using these sequences is evaluated and compared with the conventional serial search techniques.

Asymptotic Behaviour of MFSK in Noisy Phase Channels

Yeheskel Dallal and Shlomo Shamai (Shitz) *Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

The limiting performance of quadratically detected M-ary frequency shift keying (MFSK) in the presence of both additive white Gaussian noise (AWGN) and phase noise for $M \to \infty$ is addressed. A unifying analytical treatment is presented relying on power moments characterization of the prefiltered noisy phase signal. The moments are readily evaluated for a Brownian and a piecewise constant Markovian phase process. The fundamental impairments due to phase noise, random signal suppression, and crosstalk are addressed. It is shown that the error probability approaches zero with M for all information rates up to C, the maximum achievable rate. The optimum prefiltering bandwidth and the deterioration due to a finite frequency spacing are elaborated. For the case of the Brownian phase, encountered in coherent lightwave communication, the results indicate that near quantum limited performance is achievable as long as the average count of photons received during the coherence time related to the nonzero linewidth is sufficiently large. An upperbounding error rate exponent (reliability function) $E(R)$ is obtained. This exponent is a non negative decreasing convex function for all rates $0 \le R \le C$, equals zero at exactly $R = C$ and is linear for low rates $0 \le R \le R_c$ where R_c is the critical rate. The associated terms $E(R)$, R_c, C and $R_0 = E(0)$, i.e., the cut-off rate are investigated in the phase noise limited regime.

A Noncoherent CPM-Detector That Uses A Reduced Set of Basis Functions

Torgny Andersson and Arne Svensson, *Telecommunication Theory, University of Lund, Box 118, S-221 00 Lund, Sweden, and Ericsson Radar Electronics AB, Aerospace Division, Avionics and Missile Electronics, Sweden*

Bandwidth and power efficient modulation schemes has, during some years, gained in interest due to the limited frequency spectrum. Continuous Phase Modulation (CPM) is a class of modulation schemes, with a constant envelope, which can be made both very narrowbanded and power efficient simultaneously.

In some applications noncoherent detection is preferred. This detector does not have to know the phase of the carrier. The optimum noncoherent MLSE detector for CPM is, however, quite complex. In the detector the received signal is filtered in a filterbank. In this paper, the number of filters in the filterbank is reduced by using a limited signal space in the CPM-detector. This limited signal space is

found by a Gram-Schmidt procedure, where the loss in energy in each filter in the filterbank is minimized. The error performance is calculated by means of a minimum equivalent Euclidean distance. This equivalent Euclidean distance is calculated in the signal space for the mismatched detector.

It is found that the number of dimensions in the mismatched detector can be reduced quite a lot, with just a small reduction in the distance compared to the optimum detector. For a quaternary $2RC$-scheme with $h = 1/3$, the number of dimensions is reduced from 32 in the optimum detector to 6 in the mismatched detector with a loss of less than 0.5dB in the error performance.

A New Kalman Filtering Receiver over Fading Multipath Channels

P. H. G. Coelho *Rua Mar Bittencourt, 120A C/11, Riachuelo - 20951, Rio de Janeiro-RJ., Brasil*

This paper proposes a new Kalman receiver over fading multipath channels in presence of additive Gaussian noise using for channel model a rational transfer function.

The receiver has less parameters to be adjusted so its computing load is less demanding than the standard Kalman receiver.

The frequency selective fading causes a severe intersymbol interference which is harmful particularly for high capacity digital radio systems. The tracking properties of the Kalman filtering jointly with the use of a rational transfer function (with less parameters than the usual model) for the channel promote a fast convergence for the equalizer algorithm in adverse conditions such as those experienced in fading multipath channels.

Several examples are presented by means of a simulation of a digital radio communication system using a 16-QAM modulation over a frequency selective fading channel modeled by a two ray fading channel model.

The performance of the proposed equalizer, in terms of error rate, is compared with two transversal filter receivers, one using the gradient algorithm and other using the least squares method. The results indicate that the suggested receiver has a better performance particularly over the transversal filter receiver using the gradient algorithm.

Optimum Soft-Decision Demodulation for ISI Channels

S. Raghavan and G. Kaplan *Center for Magnetic Recording Research, University of California, San Diego; La Jolla, Calif. 92093, and Qualcomm, Inc., San Diego, Calif. 92121*

Two different schemes of soft-decision demodulation for channels with finite intersymbol interferance (ISI) in the presence of additive white Gaussian noise are analyzed. The first approach employs ideal interleaving; the second involves maximum-likelihood decoding (MLD) for channels with deterministic finite memory. In both schemes the cut-off rate R_0 of the discrete channel created by the soft-decision demodulator is chosen as the design criterion. Expressions for the optimal thresholds of the quantizer associated with the demodulator are derived. Results for the channel with ISI from one pulse on each side of the current pulse are presented, and the effects of ISI and the number of quantization levels are demonstrated for both schemes. The gain of the MLD approach over the interleaving scheme is shown quantitatively. Finally the (1-D) channel with soft-decision demodulation is analyzed. Closed-form solutions are derived and variation of R_0 with the number of quantization levels is presented.

Signal Processing in Channels with Intersymbol Interference

Daniel D. Klovsky *Electrotechnical Institute of Telecommunications, Kuibyshev 443071, USSR*

The paper deals with the communication system with test pulse and prediction (STPP), suggested by the author, in which the transmission of code symbols is interrupted periodically by test pulses with guard space, the duration of which is equal to $\tau = QT$ (T - transmission period, $Q = [\tau_m/T]$ - relative memory of channel, τ_m - time extent of channel memory). This system allows to realize adaptive quasi-

coherent symbol-by-symbol message deciding receiver at the time interval $T_{obs} = (1+D)\,T$ (D - decision time delay) in the channel with varying (unknown) parameters and intersymbol interference.

On the base of STPP suboptimal (when $D = 0$) symbol-by-symbol receiving algorithm using decision feedback was designed in the case of the additive Gaussian quasi-wight noise in the channel.

This STPP and Viterbi processors tree-like diagrams, error probabilities and complications are compared.

The possibility of combination of demodulation and conversion code decoding in the STPP procedure is analysed. Such a possibility is also discussed for the optimal symbol-by-symbol algorithm in the channel with intersymbol interference.

Quantization Noise Spectra

Robert M. Gray *Information Systems Laboratory, Stanford University, Stanford, California, 94305* (40 min.)

Uniform quantizers play a fundamental role in digital communications systems and have been the subject of extensive study for many decades. The inherent nonlinearity of quantizers makes analysis difficult. It usually has been accomplished either by assuming the quantizer noise to be a signal-independent, uniform white random process or by replacing the quantizer by a deterministic linear device or by combining the two assumptions. Such linearizing approximations simplify the analysis and permit the use of linear-systems techniques, but few results exist quantifying how good such approximations are for specific systems. These complications are magnified when the quantizer is inside a feedback loop, as in the case of the Sigma- Delta and Delta modulators.

Exact descriptions of the moments and spectra of quantizer noise have been developed recently for the special case of Sigma-Delta modulators. These results demonstrate that the white noise and linearization assumptions can be quite poor approximations in some systems. It turns out that many of the techniques used in the analysis were first applied to the analysis of quantizers by Clavier, Panter, and Grieg (1947) in pioneering, but often overlooked work that preceded Bennett's (1948) classic study study of quantization noise spectra.

We take advantage of the benefit of hindsight to develop several of these results in a unified and simplified manner. Exact formulas for quantizer noise spectra are developed and applied to a variety of systems and inputs, including scalar quantization (PCM), dithered PCM, Sigma-Delta modulation, dithered Sigma-Delta modulation, second-order Sigma-Delta modulation, and two-stage Sigma-Delta modulation.

SESSION WA2

MULTIPLE ACCESS II

Throughput and Delay Performance of a Channel-Sensing Coded Band-Limited Spread-Spectrum Multiple-Access Scheme

Samuel Resheff and Izhak Rubin *Electrical Engineering Department, 67310 Boelter Hall, University of California, Los Angeles, CA 90024*

Recent results indicate that wide-band nets using spread spectrum signals and operating under a CSMA access protocol (CSMA/SS) can be efficient, provided the bandwidth is wide enough so that the signal set is almost orthogonal. Signals' orthogonality eliminates packet collisions, and the CSMA/SS feature of the implementation can thus provide for better coordination of station transmissions. In particular, when half-duplex transceivers are used, a ready user will defer transmission when the channel is sensed busy, thereby unlocking its own receiver for possible message reception. However, as the number of stations in the net increases, while the total available channel bandwidth remains fixed, perfect signal orthogonality no longer prevails. Packet collisions become then the dominant factor in determining the network delay-throughput performance. Under such conditions, channel sensing can be used to provide information to stations as to when the channel is overly utilized. Such side-information can be obtained by using specially encoded message headers or by using an ancillary broadcast channel which is used solely to inform the net stations about the channel's activity level. In addition, error-correction codes are used to provide for correction of errors that occur due to simultaneous transmissions of non-orthogonal signals. In this paper, we introduce and study the message delay and channel throughput behavior of such a CSMA/SS scheme. It is assumed that each net station can gain information as to whether the total number of on-going transmissions exceeds a given threshold (M), or not. This entails, for example, the availability of an ancillary low-rate out-of-band or in-band signaling channel. A transmitting station will abort its transmission upon the reception of a signaling; message indicating that the current number of transmissions exceeds the prescribed threshold. A random-access scheme is used to control the access of messages to the channel. Using our derived performance equations, we present numerical results illustrating the delay-throughput performance of such a CSMA/SS scheme. Key parameters involved in this performance analysis include: channel bandwidth, error-correction code capability, and propagating delay. Given an average message length, an activity threshold level M can be selected to yield the best delay-throughput performance characteristics.

Capacity Region of a Waveform Gaussian Multiple-Access Channel

Chao-Ming Zeng, Ning He, and Federico Kuhlmann *Universidad Nacional Autónoma de México, Facultad de Ingenieria, División de Estudios de Posgrado, P.O. Box 70-256, 04510 México, D.F.*

A waveform (continuous-time) Gaussian multiple-access channel with "*average-power*" constrained inputs is considered. That is, the sum of two input signals $x_1(t)$ and $x_2(t)$ is passed through a linear filter $H(f)$. The input need to satisfy the average-power constraints $\frac{1}{T}\int_0^T X_i^2(t)\,dt \le P_i$, i=1, 2, for large T. The capacity region of this channel is obtained by using the orthonormal expansions of waveforms approach and a result for a parallel, discrete-time, Gaussian multiple-access channel. Our result shows that for Gaussian multiple access channels, there also exists a so-called "*water-filling*" interpretation for multi-user encoding, similarly as in the classic (single input) Gaussian channel case. Finally, an optimum assignment of the power spectra without time-sharing is given to attain the capacity line.

On Growing Random Trees in a Random Environment with Applications to Multiaccess Algorithms

Ilan Kessler and Moshe Sidi *Electrical Engineering Department, Technion - Israel Institute of Technology, Haifa 32000, Israel*

Binary trees that grow in a two-state Markovian environment are considered. For such trees we first obtain the condition on the environment process so that the tree does not grow indefinitely, and then we calculate the expected number of vertices in the tree. These two problems are addressed for two different ways of growing the binary tree, and they are compared. The two ways differ in the *order* by which the vertices of the growing tree reproduce.

The above results are applied to multiaccess networks in which the shared channel is noisy. We assume a slotted-time collision-type channel, Poisson infinite-user model, and a binary feedback. Due to the noise in the shared channel the received signal may be detected as a collision even though no message or a single message is transmitted (referred to as *error*). A common assumption in all previous studies of multiaccess algorithms in channels with errors is that the channel is *memoryless*. The above results are applied to the analysis of the operation of the tree-CRA with *memory*. We obtain the condition on the channel parameters for stability and the throughput of the algorithm when this condition holds.

Polling Systems with Routed Customers

Moshe Sidi and Hanoch Levy *Electrical Engineering Department, Technion - Israel Institute of Technology, Haifa 32000, Israel, and Computer Science Department, Tel-Aviv University, Tel-Aviv 69978, Israel*

A queueing network that is served by a single server in a cyclic order is studied. Customers arrive at the queues from outside the network according to independent Poisson processes. Upon completion of his service, a customer may leave the network, be routed to another queue in the network or rejoin the same queue for another portion of service. The single server moves along the different queues of the network in a cyclic manner. Whenever the server arrives at a queue (polls the queue), he serves the waiting customers in that queue according to some service discipline. Both the gated and the exhaustive service disciplines are considered. The service time of a customer has a general distribution (may be different from queue to queue). When moving from one queue to the next queue, the server incurs a switch-over period with a general distribution.

For this general queueing network we derive the expected number of customers present in the network queues at arbitrary epochs, and compute the expected delays observed by the customers. In addition, we introduce a pseudo conservation law for this network of queues. Some interesting numerical examples illustrate how routes and server moving direction affect the performance of the network.

On Gaussian Feedback Capacity

Amir Dembo *Information Systems Laboratory, Stanford University, Stanford, CA 94305*

Pinsker and Ebert proved that in channels with additive Gaussian noise, feedback at most doubles the capacity. Recently, Cover and Pombra proved that feedback at most adds half a bit per transmission. Upper bounds on the feedback capacity are obtained here using majorization tools. These bounds imply for noise covariance matrices with bounded away from zero and infinity eigenvalues that at high input SNR the difference between feedback and non-feedback capacities is marginal, while at low input SNR their ratio is close to one. Specializing these results to stationary channels we recover some of the bounds recently obtained by Ozarow using different techniques.

Modified Viterbi Decoding for Convolutionally Encoded Hybrid-ARQ Protocols

Stephen B. Wicker and Bruce Harvey *School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332*

An error pattern in a convolutionally encoded data packet corresponds to a cycle containing the zero state in a weighted directed graph. The structure inherent in this error process can be used to define a type-I hybrid-ARQ protocol. Let the "transition trap length" J_T be the minimum number of branches a path leaving the zero state must take before its weight is guaranteed to equal or exceed that of the minimum free Hamming distance d_{free}. Note that the path need not terminate at the zero state. In the proposed modification to Viterbi decoding the rate of increase of the partial path metrics is monitored over a sliding window of J_T branches. A path is labeled unreliable if the rate of increase of its partial path metric exceeds a threshold τ. If the maximum likelihood path is found to be unreliable, then a retransmission of the packet is requested; otherwise, decoding is allowed to proceed normally. Analysis and simulation results are presented to show that the proposed type-I hybrid-ARQ protocol substantially improves data reliability at the expense of a minimal reduction throughput.

Coding Theory for Secret Sharing Communication Systems with Two Gaussian Wiretap Channels

Hirosuke Yamamoto *Department of Communications and Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan*

The coding theorem is proved for the secret sharing communication system (SSCS) with two Gaussian wiretap channels, which is an extension of both the SSCS with two noiseless channels and the Gaussian wiretap channel (GWC) system. The admissible region of rates and security levels for the SSCS with two GWCS's is given by

$$\begin{cases} h_1 & \le C_{S_1} R_1 + C_{M_2} R_2 \\ h_2 & \le C_{M_1} R_1 + C_{S_2} R_2 \\ H(S) & \le C_{M_1} R_1 + C_{M_2} R_2 , \end{cases}$$

where C_{M_j} and C_{S_j} $(j = 1, 2)$ are the channel capacity and secrecy capacity of GWC j, respectively, rate R_j is defined by channel-j symbols per source symbol, and security level h_j is measured by the equivocation of wiretapper j.

On the Characterization of Information Divergence

G. Q. Shi *Telecom Research Laboratories, 761-772 Blackburn Road, Clayton, Victoria 3168, Australia*

This paper presents some results on single letter characterization of information divergence.

For the discrete case, information divergence is degenerate as data are compressed, i.e., when one encodes the data in a composition class, the information, at the level of information divergence, is not, in fact, compressed. Therefore, by using a normal definition of coding rate, one cannot obtain characterization of the achievable single-letter solution for the information divergence.

Here we first give a bent rate definition and then, by using it, we get some results for one variable and bivariables with data compression. For the problem of bivariables with two-side data compression, however, only bounds of the single letter characterization are obtained.

SESSION WA3

SIGNAL PROCESSING II

On the Estimation of the Order of a Stationary Ergodic Markov Source

Chuangchun Liu *Electrical Engineering Department, University of Maryland, College Park, MD 20742*

We study an estimator for the order of a stationary ergodic Markov source. This estimator, which is easy to implement, is consistent and possesses exponential and near-exponential rates of decay of the probabilities of underestimation and overestimation respectively.

Efficient Identification of Impulsive Channels

Serena M. Zabin and H. Vincent Poor *School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, and Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801*

The man-made electromagnetic environment, and much of the natural one, is basically impulsive, i.e., it has a highly structured form characterized by significant probabilities of large interference levels. It has been demonstrated that the performance of communications, radar, and sonar systems operating in impulsive channels can be greatly enhanced if the statistics of the channel are known and exploited. Consequently, the problem of identifying impulsive noise channels is a basic and important one. A physically-meaningful, parametric model for impulsive interference is the so-called Class A Middleton model, whose parameters A and Γ can be adjusted to fit a great variety of non-Gaussian noise phenomena occurring in practice. The first parameter, A, referred to as the "Overlap Index," is a measure of the average overlap of successive emission events. The second parameter, Γ, the "Gaussian Factor," is the ratio of the intensity of the independent Gaussian component of the input interference to the intensity of the non-Gaussian component.

In this study, a batch estimator of the Class A parameters with good small-sample-size performance is obtained. This estimator is based on the EM algorithm, a two-step iterative technique which is ideally suited for the Class A estimation problem since the observations can be readily treated as "incomplete data." For the single-parameter estimation problem (A unknown, Γ known), a closed-form expression for the estimator is obtained. Furthermore, for the single-parameter estimation problem, it is shown that the sequence of estimates obtained via the EM algorithm converges, and if the limit point to which the sequence converges is an interior point of the parameter set of interest, then it must necessarily be a stationary point of the traditional likelihood function. The small-sample-size performance of the proposed EM estimator is also examined via an extensive simulation study. For both the single-parameter and two-parameter estimation problems, the results of this study indicate that this likelihood-based scheme yields excellent estimates of the Class A parameters (in terms of attaining the Cramér-Rao Lower Bound) for small sample sizes ($O(10^2)$).

Nonparametric Identification of a Cascade Nonlinear Time Series System

Miroslaw Pawlak *Department of Electrical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2*

In this paper a nonlinear time series system of a cascade structure is identified. The system consists of a nonlinear memoryless element followed by a dynamic linear subsystem. Given a colored Gaussian input, the Hermite polynomials based method for the recovering of the system nonlinearity is presented. The algorithm is carried out by solving an integral equation related to the identification problem. The consistency along with rates of convergence are established. No assumptions concerning continuity of the characteristic or its functional form are required. An extension to the case of an input Markov process possessing the bivariate density with a diagonal expansion is also made.

Application Criteria of the Pencil of Functions Method in ARMA System Identification

Diamantino R. S. Freitas *Faculdade de Engenharia da Universidade do Porto, D.E.E.C., 4099 Porto Codex, Portugal*

ARMA system identification is generally considered to be a difficult nonlinear problem. the use of the Kalman equation error, as in the pencil of functions method (POFM), turns that into a linear alternative, yielding good results in low-noise situations. the identification of electroacoustic transfer functions, in high accuracy, low noise, conditions, is a domain well suited for the application of the POFM, given its robustness, consistence and low bias. However, the success in the application and the accuracy of the results depend, in a crucial way, on the appropriate selection of **q**, the pole of cascaded first-order low pass filters used in the method and **n**, the number of signal samples employed. This essential point has been, to our knowledge, paid little attention. The work presented shows that the error energy **E** presents a sharp minimum. this leads to the formulation of criteria: the optimum **q** value if close below the lowest of the system poles amplitudes; the optimum **n** value is a compromise between increasing errors in the estimated model numerator and decreasing errors in the denominator.

Approximate Bayesian Classification Based Upon Hidden Markov Modeling

Neri Merhav and Yariv Ephraim *Speech Research Department, AT&T Bell Laboratories, Murray Hill, NJ 07974*

We investigate a Bayesian approach to multiple hypothesis testing for hidden Markov sources, whose statistics are given empirically by training data. The exact Bayesian optimal decision rule involves calculation of conditional means of probability density functions and hence it is computationally untractable. To avoid this difficulty, we propose an alternative decision rule which is computationally more attractive. It is proved that the asymptotic exponential rate of decay of the error probability, associated with the proposed decision rule, is optimal. Furthermore, it does not require knowledge of the prior probability densities of the model parameter. The approach is generalized to hypotheses testing in a noisy environment, given training data of the clean sources and the noise process. Simulation results on computer generated hidden Markov processes reveal significant preference of the proposed approach over a standard method currently used.

Imaging a Randomly Translating Object from Point Process Observations

Donald L. Snyder and Timothy J. Schultz *Electronic Systems and Signals Research Laboratory, Department of Electrical Engineering, Washington University, St. Louis, MI 63130*

Measurements of a randomly translating object described by a spatial intensity function are modeled as a time-space doubly stochastic Poisson process in which the intensity function moves randomly in time. A method for producing constrained maximum-likelihood estimates of the object intensity is developed and examples for various characterizations of the motion are presented.

Robust Signal Reconstruction in a Hilbert Space Setting

Richard J. Barton *Orincon Corporation, 9363 Towne Centre Drive, San Diego, CA 92121*

In this presentation, we examine the problem of reconstructing an unknown signal from noisy observations. We assume throughout that the signal is a member of a known *reproducing kernel Hilbert space* (RKHS) H so that the observation functionals $\mathbf{L}_t(f) = f(t)$ are bounded. We assume also that the true values of the signal f_0 are known only to belong to some convex set C containing the observations $\{y(t), t \in \mathsf{O}\}$. In order to generate a robust reconstruction, we seek a function $\hat{f} \in \mathsf{H}$ that solves the *minimax* problem

$$\sup_{f \in \mathsf{U}} \|f - \hat{f}\|_{\mathsf{H}}^2 = \inf_{g \in \mathsf{H}} \sup_{f \in \mathsf{U}} \|f - g\|_{\mathsf{H}}^2 ,$$

where $U = \{f \in H : \|f\|_H \leq M, \{f(t), t \in O\} \in C\}$, and M is an arbitrary constant. We show that $\hat{f}$ is simply the minimum norm element of U closure, and we give some examples for which we can characterize $\hat{f}$ explicitly.

Asymptotics of Divergent LMS

Todd F. Brennan *Department of Electrical and Computer Engineering, University of Wisconsin, 1415 Johnson Drive, Madison, WI 53706*

Convergence properties of the LMS algorithm have been well studied [2,8,9,10], but few quantitative results are available for divergent LMS. Large deviation theory is used to compute several different LMS asymptotics. During the derivation, it is observed that LMS does not scale as a slow Markov walk [1], so a transformed process is used instead. Necessary and sufficient conditions for the existence of large deviation behavior in LMS are given. For certain inputs, it is proven that LMS diverges exponentially with probability one, and mean exit time can be computed in closed form. Convergence is shown using large deviation theory as well. Direct simulations confirm theoretical predictions. The method shown here is general, and can be used to handle the observation noise case and others. Most notably, these large deviation methods handle non-Gaussian inputs and Markov input dependencies, but may not admit closed form solutions for certain data dependencies.

Double Sampling *M*-Detection Procedure

Liu Youheng and Tang Chuanzhang *Department of Radio Electronics, Peking University, Beijing, P.R. of China.*

A new detection procedure, the double sampling procedure, is developed in this paper. Applying it to the detection of signals in contaminated Gaussian noise, we obtain the double sampling *M*-detector. It exhibits simplicity in structure while retaining the robustness property of the *M*-sequential detector. Simulation results are given that coincide with the analytical results quite well.

SESSION WA4

SOURCE CODING II

Entropy-Based Bounds on the Redundancy of Prefix Codes

Padhraic Smyth *Communication Systems Research, Jet Propulsion Laboratories 238-420, 4800 Oak Grove Drive, Pasadena, CA 91109*

We consider the problem of bounding the redundancy of binary prefix codes for discrete memoryless sources. Recent results have bounded the redundancy of Huffman codes in terms of the probability of various components of the source alphabet, e.g., the most likely letter. Similar results for alphabetic prefix codes have also been derived where the smallest probability component is known. We extend these results to the case where the entropy of the source can be estimated or bounded, and derive a variety of bounds for the redundancy of Huffman, Shannon-Fano, and weight-balanced (alphabetic) coding schemes, expressed in terms of source entropy.

Efficient Representations for Huffman Coding

Cheng-Chang Lu *Department of Mathematical Sciences, Kent State University, Kent, OH 44242*

An efficient representation for the Huffman tree is proposed. The tree is uniquely specified by a sequence of decimal node numbers. The Huffman code can be determined from the node number easily by a shifting procedure, and the node number can also be generated from the corresponding code by adding a leading bit. A new data structure for static Huffman coding is developed using this representation. Based on the inherent parent-child property, an efficient dynamic Huffman coding algorithm can also be implemented without building a look-up table to keep track of all parents and children.

A Universal Model Based on Minimax Average Divergence

Cheng-Chang Lu and James George Dunham *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, and Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275*

For a non-adaptive data compression system, the source structure and statistics required by the coding unit have to be determined from the input sequence before the corresponding codes can be generated. This requires each individual input sequence to pass the source modeling part of the system first and then go through the coding part. Such implementation may not be acceptable in terms of extra memory and time required. In this paper, a universal model for a class of input sequences is proposed, which minimizes the maximum average divergence between the model and training samples. A theoretical searching algorithm for finding the minimax average divergence is developed, based on the property that the maximum average divergence can be decreased by modifying the model. Also proposed is a practical searching algorithm that can be easily implemented on digital computer to find the minimax universal model.

A New Asymptotically Optimal Code of the Positive Integers

Hirosuke Yamamoto and Hiroshi Ochi *Department of Communications and Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan*

A new universal binary code of the positive integers is proposed as a modified version of Wang's flag encoding scheme based on the bit stuffing technique. The average codeword length defined by $\sum_{n=2^M}^{2^{M+1}-1} L(n)/2^M$, where $L(n)$ is the codeword length of n, is one bit shorter than Wang's scheme if the flag length f is two, and it is shorter than Wang's if $M \leq 12$ for $f = 3$, $M \leq 21$ for $f = 4$, and so on. The

performance of the new scheme is also compared with other universal schemes for geometrically or Poisson distributed integers. The new scheme is almost as efficient as Capocelli's Fibonacci encoding scheme and is more efficient than other schemes. Moreover, the encoding and decoding algorithm is simpler than other universal schemes including Wang's. Hence, the new flag scheme is suited for hardware implementation as well as software in many practical applications. Furthermore, an asymptotically optimal code can be realized by modifying the new flag scheme such that the flag length varies dynamically.

Combined Equalization and Coding with Minimum Mean Square Vector Coding

J. M. Cioffi, J. S. Chow, and J. Tu *Information Systems Laboratory, Stanford, CA 94305*

The combined design of equalization and coding methods is necessary to achieve the highest possible data transmission rates on channels with intersymbol interference. Existing approaches to combined equalization and coding are strictly zero-forcing in that intersymbol interference is eliminated through feedback in symbol-by-symbol approaches or through careful selection of successively transmitted blocks in block ("multitone" or "vector coding") approaches. In this paper, we investigate block approaches where some residual intersymbol interference is permitted in order to achieve an overall reduction in mean-square distortion (noise plus residual interference).

The approach involves converting the ISI-channel into a canonical minimum-mean-squared-error minimum-phase equivalent through the use of what is called a "mean-square whitened matched filter." this conversion leaves the overall system with the highest possible SNR. Further, the optimum applied transmit spectrum to such an equivalent channel is shown to be the "water-pouring" energy distribution. We then design a coset-coded multichannel modulation method with this spectrum for the equivalent mean-square-minimum-phase channel and apply it to the channel. We call this method "Minimum Mean Square Vector Coding," or MMSVC. MMSVC is shown to achieve transmission rates that are as close to capacity on the ISI-channel, as the equivalent coset code is to capacity of the "flat" (or "ideal") additive white Gaussian noise channel (2-4dB away from capacity), even at low to moderate SNR's where existing zero-forcing methods do not apply as well.

Bounds to the Capacity of Discrete Memoryless Channels with Input Constraints

Ali Khayrallah and David L. Neuhoff *Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109*

Bounds are presented to the capacity of a discrete memoryless channel with input constraints. Specifically, Mrs. Gerber's Lemma is used to derive a simple lower bound to the capacity of a binary symmetric channel with the constraint that the infinite input sequence be a member of a specified subshift. this bound is a function only of the crossover probability of the channel and the capacity of the subshift. Secondly, by restricting the allowable input probability distributions (in the usual definition of *nth*-order capacity), we obtain upper bounds to the capacity of a discrete memoryless channel with inputs constrained by a subshift. One such restriction is the requirement that the n-dimensional input distribution be stationary. Others derive from the maximum and minimum frequencies with which strings can occur in a member of the subshift. The bounds are evaluated for binary symmetric channels and (d,k) run-length constraints. Comparisons are made with the recent bounds of Zehavi and Wolf.

The Adaptive Guazzo Algorithm

Gérard Battail *Ecole Nationale Supérieure des Télécommunications, Département Communications, 46 Rue Barrault, 75634 Paris Cedex 13, France*

The Guazzo algorithm is a source coding algorithm which can easily be made adaptive in order to deal with unknown and/or varying source statistics. This paper describes the adaptive version of this algorithm, reports theoretical predictions of its performance and results of its simulation on both stationary and nonstationary sources, including actual computer files.

On the Optimal Inductive Inference Scheme from the View Point of Source Coding

Toshiyasu Matsushima, Joe Suzuki, Hiroshige Inazumi, and Shigeichi Hirasawa *Dept. of Management Information, Yokohama College of Commerce, Yokohama 230, Japan; Dept. of Industrial Engineering and Management, Waseda University, Tokyo 169, Japan; and Dept. of Information Science, Sagami Institute of Technology, Kanagawa 255, Japan*

In this paper, we discuss the inference of predicate logic, which is widely used for representation of knowledge in artificial intelligence (AI) systems, from the view points of source coding and decision theory. Since the inference in logic can be regarded as some kinds of information transformation, we can recognize an analogy between inference and source coding. Inductive inference is regarded as source encoding, because observed facts or examples are compressed into an axiom similar to a source sequence into a codeword. On the contrary, deductive inference is interpreted as decoding. >From the view point of decision theory, inductive inference is regarded as the decision problem selecting the collect axiom which represents an observing world. We propose a new inductive inference scheme which induces the minimum Bayes risk. Moreover, we show the method for selecting the axiom which represent the finite observed facts by the minimum description length code.

SESSION WA5

SHANNON THEORY III

Successive Refinement of Information

William H. Equitz *Information Systems Laboratory of the Department of Electrical Engineering, Stanford University, Stanford, CA 94305; now at IBM Almaden Research Center, San Jose, CA 95120-6099*

We characterize problems in which optimal descriptions can be considered as refinements of one another. We do this because we may optimally describe a message with a particular amount of distortion and later decide that the message needs to be specified more accurately. If an addendum to the original message is then sent we hope that this refinement is as efficient as if the more strict requirements had been known at the start. In general, we ask whether it is possible to interrupt a transmission at any time without loss of optimality.

We present necessary and sufficient conditions for achieving optimal successive refinement and establish that all finite alphabet problems have the required properties for at least small distortions. Furthermore, we show that finite alphabet signals with Hamming distortion, Gaussian signals with squared error distortion, and Laplacian signals with absolute error distortion all satisfy these requirements over the entire range of possible distortions. On the other hand, we exhibit a family of simple counterexamples which show that successive refinement is not achievable in general.

Maximum Entropy Charge Constrained Run Length Codes

Kenneth J. Kerpez, Ayis Gallopoulos, and Chris Heegard *Bell Communications Research, 445 South Street, Morristown, NJ 07960, Athens, Greece, and the Department of Electrical Engineering, Cornell University, Ithaca, NY 14853*

The maximum entropy distribution maximizes the code rate for a given channel constraint. Closed form expressions are known for the power spectrum of maximum entropy run length limited (d,k) sequences and charge constrained (C) sequences. Information is coded into a sequence that satisfies both run length and charge constraints simultaneously; the charge constrained run length limited (d,k,C) sequence. A DC null is implied by the charge constraint C, making it useful for a recording system with an AC coupled rotary head. An expressing for computing the maximum entropy distribution and its power spectrum is presented for a (d,k,C) sequence. The expression involves the adjacency matrix of the variable length state transition diagram of the (d,k,C) sequence. Simplified formulas are given for the case of no K constraint, and for the tight constraint $k = d + 1$.

Gambling Using a Finite-State Machine

Meir Feder *12 Shirat-Hazamir St., Herzliya, 46420, Israel*

Sequential gambling schemes in which the amount wagered on the outcome of a random sequence is determined by a finite state (FS) machine are defined and analyzed. We assume that the FS machine determines the fraction of the capital wagered at each time instance, i, on the outcome at the next time instance, $i + 1$, and that wagers are paid at even odds. We show that the maximal capital gain in any finite state sequential gambling scheme is given by,

$$S_n = S_0 2^{n\left[1 - H^{FS}(\underline{x})\right]}$$

where S_0 is the initial capital, S_n is the capital after time instance n, $\underline{x}$ is the outcome sequence and $H^{FS}(\underline{x})$, which is a measure based on the empirical entropy of $\underline{x}$, is defined as the finite-state complexity of $\underline{x}$. We also analyze a specific gambling scheme based on the Ziv-Lempel method for universal

compression, and its performance provides a relation between the finite-state complexity defined above and the finite-state complexity defined by Ziv and Lempel.

An Application of the Galileo Multidimensional Scaling System to Human Communication

Walton B. Bishop *University of Maryland, College Park, MD 20742*

The multidimensionality of effects that a message recipient's prior knowledge has upon comprehension calls for new methods of analysis. Attempts to reduce multiple interpretations of a given message by using causal analysis proved to be unduly cumbersome. The Galileo multidimensional scaling system, however, provides a simple, yet precise, way of measuring some of the effects prior knowledge has on message interpretation. Data collected from a representative sample of message recipients, when analyzed by the Galileo system, will tell the message originator how to modify a message so that it will be interpreted correctly by its intended audience. The Galileo™ computer program has been used successfully in such diverse areas as mass communication, political communication, criminal justice, advertising, and marketing. It seems ideally suited for the unobtrusive measurement of the effects produced by a message recipient's prior knowledge. It is these effects that usually interfere with the application of Shannon theory to human communication. (The research reported here was done as part of the author's doctoral dissertation under the direction of Professor Edward L. Fink, Department of Communication Arts and Theatre.)

On Practical Applications of the ITRULE Algorithm

Rodney M. Goodman and Padhraic Smyth *Department of Electrical Engineering, Caltech 116-81, Pasadena, CA 91125, and Communication Systems Research, JPL 238-420, Pasadena, CA 91109*

In a previous paper we described the ITRULE learning algorithm which derives the most informative set of probabilistic rules from a set of sample data. The algorithm uses information-theoretic bounds to optimally constrain its search in the exponentially large space of possible rules. Here we describe our most recent work on the algorithm. We define the classes of problems which require such general sets of rules rather than more specific solutions such as decision trees. Typically these problems involve partial and missing information, initial context information, and require joint probability estimates over arbitrary subsets of domain variables. We describe the application of ITRULE to classification problems of this nature, using real data such as medical databases and congressional voting records. We also describe how ITRULE can be used as a knowledge acquisition tool to extract rules from databases where no domain experts exist. In particular, we describe how the algorithm can infer fault diagnosis and situation assessment rules via causal simulations of complex man-made systems which have not yet been built, for example, subsystems of NASA's planned space station and JPL's Mars rover.

A Dual Control Strategy to Minimize the Discrimination Information for Stochastic Systems

Charles D. Schaper, Duncan A. Mellichamp, and Dale E. Seborg *Department of Chemical and Nuclear Engineering, University of California, Santa Barbara, CA 93106*

A new dual control policy is synthesized to minimize the discrimination information associated with a stochastic system and a nonstationary Gaussian-distributed performance specification. The system is characterized by a linear AR model with stochastic coefficients. The first and second moments of these stochastic coefficients are estimated recursively by a Kalman Filter. To indicate the presence of a dual effect, the mutual information measure is employed to quantify the statistical relationship between the estimated AR model coefficients and the available data. A parametric analysis of the mutual information measure is conducted to prove that the new stochastic control strategy possesses an inherent dual effect in terms of the fundamental parameters of the control system synthesis problem; whereas previous dual control strategies have had to incorporate an *ad hoc* weighting coefficient to obtain the dual effect. The relationship of the proposed dual control strategy with suboptimal dual policies is established via

theoretical analysis and simulation. Hence, an important link between information theory and systems theory is established through this new information theoretic control strategy for stochastic systems.

On the Autocorrelation Functions of Binary Sequences Obtained from Finite Geometries

Agnes Hui Chan, Andrew Klapper, and Mark Goresky *College of Computer Science, Northeastern University, 360 Huntington Ave., Boston, MA 02115*

Maximum period linear feedback shift register sequences with non-linear feedforward functions have been used in modern communication systems. Many of these sequences are required to have high linear complexities and good autocorrelation function values. Recently, Chan and Games introduced a class of binary sequences obtained from finite geometries using nonlinear feedforward function $p : GF(q) \rightarrow GF(2)$, with q odd. They showed that these sequences have high linear complexities. Brynielsson had studied similar problem with q even and established the linear complexities of these sequences in terms of the polynomial expression of the function p. In this paper, we consider the autocorrelation functions of these sequences, and established their values in terms of the autocorrelation values of the sequence obtained from $(p(\beta^0), p(\beta), \cdots, p(\beta^{q-2}))$, where β is a primitive element of $GF(q)$.

Digital Synchronous Processes Generated by a Stationary and Independent Symbol Sequence--General Properties

Adolfo V. T. Cartaxo and Augusto A. de Albuquerque *Instituto Superior Técnico (IST), DEEC/CAPS, Av. Rovisco Pais, P-1096 Lisboa, Portugal*

We study the digital synchronous processes (DSP) generated by a stationary and independent symbol sequence. The contributions of this paper are: i) the calculation of a general expression for the nth order moment generating function (MGF) and the conclusion that these processes are strict-sense cyclostationary (SSCS); ii) from the nth order MGF, the calculation of general expressions for the expected value, autocovariance, autocorrelation and power spectral density (PSD) of these processes; iii) from this theory, the computation of the PSD of a digital pulse position modulation (DPPM) system.

Existence, Construction Methods and Enumeration of Higher Dimensional Hadamard Matrices

Yang Yi Xian *PO Box 145, Dept. of Inform. Eng., Beijing University of Posts and Telecomm. P.R. China*

The following open problems (appeared in Trans. IEEE IT-25, 566-572, 1979) are solved in this paper: (a): A few conjectures are proved. (b): many new construction methods are shown. (c): the upper and lower bounds for the enumeration are proposed. (d): there exist no absolutely improper n-dimensional Hadamard matrices of order 2. (e): there are 4128 4-dimensional Hadamard matrices of order 2.

SESSION WA6

CODING THEORY V

Decoding Cyclic and BCH Codes up to the Hartmann-Tzeng and Roos Bounds

G. L. Feng and K. K. Tzeng *Department of Computer Science and Electrical Engineering, Lehigh University, Bethlehem, PA 18015* (40 min.)

There are many cyclic and BCH codes having their minimum distance lower bounded by the Hartmann-Tzeng (HT) or the Roos bound. In decoding such codes, multiple syndrome sequences are available. In this paper, the application of multisequence shift-register synthesis algorithms - a generalized euclidean algorithm and a generalized Berlekamp-Massey algorithm - to decode such codes up to the HT bound and the Ros bound is formally considered. The main task is to determine under what condition or conditions the connection polynomial of the shortest linear feedback shift-register obtained by the algorithms will be the error-locator polynomial. We give a detailed treatment on the application of the generalized Euclidean algorithm only since the application of the generalized Berlekamp-Massey algorithm follows similarly. For decoding up to the HT bound, we have shown that the algorithms will always produce the error-locator polynomial. However, for decoding up to the Roos bound, this is not the case unless additional conditions are satisfied. For cases that these conditions are not satisfied, the syndrome sequences are shown to be linearly dependent. Based on this dependence relation, an alternative decoding procedure is derived.

Pseudocyclic (n,k) MDS Codes over $GF(q)$

Arvind Krishna and Dilip V. Sarwate *Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana IL 61801*

We consider the existence of nontrivial pseudocyclic (n,k) maximum-distance-separable (MDS) codes modulo $(x^n - a)$ over $GF(q)$. When n is a divisor of $q+1$, such codes exist for all k if n is odd, but if n is even, such codes exist only for odd k (only for even k) whenever a is (is not) a quadratic residue in $GF(q)$. When n is a divisor of $q-1$, pseudocyclic MDS codes exist if and only if the multiplicative order of a is a divisor of $(q-1)/n$. (This research was supported by the U.S. Army Research Office under contracts DAAG29-84-K-0088 and DAAL03-87-K-0097.)

Quasi-Cyclic Codes on the Klein Quartic over $GF(2^r)$: A Procedure for Correcting 1 or 2 Errors.

A. Thiong-Ly *Dept. of Mathematics, University of Toulouse Le Mirail, 5, Allées Antonio Machado, 31058 Toulouse, France*

Let $r \equiv 0 \bmod 3$. We construct a class of quasi-cyclic codes over $GF(2^r)$ derived from the Klein quartic: $X^3Y + Y^3Z + Z^3X = 0$. An easy procedure for correcting one or two errors is proposed, which needs at most 200 products in the finite field $GF(2^r)$.

Generalized Remainder Decoding Algorithm for Reed-Solomon Codes

Masakatu Morii and Masao Kasahara *Department of Electronics and Information Science, Kyoto Institute of Technology, Matsugasaki, Sakyoku, Kyoto, 606 Japan*

Investigating efficient decoding algorithms of Reed-Solomon codes, whose minimum distance is large, is very important from both theoretical and practical points of view. In 1982, E. R. Berlekamp and L. Welch proposed a new decoding algorithm for RS-BCH codes without computing the syndromes. It is called *remainder decoding algorithm*. It is very interesting that its *key-equation* is quite different from that of conventional algorithms; for example, Peterson algorithm, Berlekamp-Massey algorithm and Euclidean (Sugiyama-Kasahara-Hirasawa-Namekawa, SKHN) algorithm.

In this paper we shall present another type of the remainder-decoding algorithm. Furthermore we shall give very useful properties of our generalized algorithm for decoding Reed-Solomon codes as fast as possible.

Systematic Decoding of Reed-Solomon Codes

Ron M. Roth and Abraham Lempel *Department of Computer Science, Technion - Israel Institute of Technology, Haifa 32000, Israel*

An $r \times r$ matrix $A = [a_{ij}]_{i,j=0}^{r-1}$ is called circulant if $a_{ij} = a_{0,j-i}$ (indices taken modulo r). In many cases, we can transform the parity-check matrix of an $[n,n-r]$ Reed-Solomon (RS) code into $H = [A^{t-1} A^{t-2} \cdots A I]$, forming a concatenation of several circulant matrices. A decoding procedure for RS codes is derived, based on such a representation of the parity-check matrix. The key idea in the proposed algorithm is a transformation of the Berlekamp-Massey algorithm into the time-domain using an r-dimensional inverse Fourier transform, compared with the n-dimensional transform used in Blahut's decoder. The decoding algorithm consists of the following steps: (i) syndrome evaluation, using the encoding circuit; (ii) finding the error-locator and error-evaluator polynomials using a reduced version of Blahut's decoder; and (iii) interpolation of the outputs of Step (ii), reusing the encoding circuit. The resulting procedure inherits both the (relatively low) time complexity of the Berlekamp-Massey algorithm, and the hardware simplicity characteristic of Blahut's time-domain algorithm. In particular, the required memory size is proportional to r, rather to n.

The Cannibalistic Traits of Reed-Solomon Codes

Oliver Collins *Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218*

This paper presents a new kind of algebraic code formed by combining Reed-Solomon codes with themselves. The concatenation scheme allows longer code words than Reed-Solomon codes with the same symbol set. Examples of the performance gains with a constraint length 15, maximum likelihood decoded, convolutional inner code appear at the end. Although the goal of coming as close as possible to channel capacity on the Gaussian channel drove the design of the codes presented, they are also suitable for the photon channel and, to a lesser extent, the magnetic recording channel.

Decoding of Reed-Solomon Codes Using Bit Level Soft Decision Information

Alexander Vardy and Yair Be'ery *Department of Electronic Communications, Control and Computer Systems, Tel Aviv University, Ramat Aviv 69978, Tel Aviv, Israel*

In this paper we present a Reed-Solomon decoder that makes use of bit soft-decision information. A Reed-Solomon generator matrix which possesses a certain inherent structure in $GF(2)$ is derived. This structure enables representation of the code as a union of cosets, each coset being an interleaver of several binary BCH codes. Such partition into cosets provides a clue for efficient bit level soft decision decoding. The proposed decoding algorithms are several orders of magnitude more efficient than conventional techniques for many codes.

SESSION WA7

TRELLIS CODING III

Noise Effects on M-ary PSK Trellis Codes

Gideon Kaplan and Ephraim Zehavi *Ministry of Defence, Tel-Aviv, Israel and Department of Electrical Engineering, Technion, Israel Institute of Technology, Technion City, Haifa, 3200, Israel, and Qualcomm, Inc, 10555 Sorrento Valley Road, San Diego, CA 92121*

In this work we present upper bounds on the error performance of M-ary PSK trellis coded system over the AWGN and in the presence of Tikhonov distributed phase noise. this model is applicable to a first-order phase tracking loop perturbed by thermal and carrier phase noise.

Our model is as follows. A binary sequence at the transmitter is encoded using a rate $R=(n-1)/n$ trellis code having an S-state encoder. The n-tuple produces one of $M=2^n$ PSK signals. The channel produces at the output a noisy discrete-time sequence $\{y_p rcurl, y_p=\rho_p x_p+n_p$. Here, $\rho_p=\exp(j\theta_p)$ is a unit vector possessing a Tikhonov distribution. It is assumed that the phase noise process is narrowband with respect to the data rate. An ideal interleaver/deinterleaver system is also assumed. We are interested in the error performance of this coded system.

Using Chernoff bounding techniques the pairwise error probability of the decoder, which uses the squared Euclidean distance as its metric, is estimated. This bound is then used to calculate the "generalized Ro". the degradation due to phase noise is demonstrated for 8-PSK and 16-PSK.

The error performance of some practical coding methods for 8-PSK and 16-PSK are also presented, based on the modified generating function approach. It is shown that there exists some codes which are more robust for a given level of phase noise, and that these codes have parallel transitions in their trellis, thus their minimum Euclidean distance is not necessarily optimal. However, they exhibit a greater degree of robustness to phase noise than the optimal codes for the AWGN channel.

Bidirectional Trellis Decoding

Farhad Hemmati *Comsat Laboratories, 22300 Comsat Drive, Clarksburg, MD 20871*

Bidirectional trellis decoding is a reduced-complexity method for soft-decision decoding of block codes. Unlike the Viterbi algorithm, which processes every trellis state, the bidirectional decoding algorithm takes advantage of the structure of the considered code to identify and process a small subset of paths in the trellis diagram containing the most likely path. The algorithm examines a received block of channel symbols in the forward and in the backward direction and selects the most likely codeword as the transmitted message. The bidirectional algorithm takes significantly fewer computations per decoded bit. For example, it takes 647 binary operations for decoding the extended Golay code, whereas Forney's decoding method requires 1351 binary operations. Extensive analysis and computer simulations indicated that for special classes of block codes, including the extended Golay code, the BER performance of the bidirectional decoding algorithm is equivalent to that with a maximum-likelihood decoder.

Rotationally Invariant Trellis Codes

Steven S. Pietrobon, Daniel J. Costello, Jr., and Gottfried Ungerboeck *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556, and IBM Research Division, Zurich Research Laboratory, Saumerstrasse 4, CH-8803 Ruschlikon, Switzerland*

A general parity check equation for rotationally invariant rate $k/(k+1)$ trellis codes is presented. This equation is different from the traditional linear parity check equations (which do not give rotationally invariant codes) in that modulo M $(M>2)$ arithmetic is used to derive the parity check equation. The signal set mapping also needs to be related to modulo-M arithmetic in terms of its rotational properties (e.g., MPSK and other "naturally mapped" signal sets).

The final parity check equation still uses modulo-2 addition for the binary sequences, but non-linear terms are added into the equation. These non-linear terms overcome the effect of a phase rotation. In contrast to the linear codes, the encoders derived from the non-linear parity check equation may not be minimal.

Using this parity check equation, rotationally invariant rate 2/3 trellis codes with 8PSK modulation are presented. These codes exhibit a coding gain slightly inferior to linear codes, but they have more tolerance to phase slips within the demodulator. (This work was supported by NASA grant NAG5-557 and by NSF grant NCR89-03429.)

Trellis Coding using Multi-Dimensional QAM Signal Sets

Steven S. Pietrobon and Daniel J. Costello, Jr. *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556*

A method of finding good trellis codes with multi-dimensional (multi-D) QAM modulation is presented. Using the 16QAM signal set, 4-D, 6-D, and 8-D QAM signal sets are constructed which have good partition and phase rotational properties.

The good partition properties are achieved by the use of block codes and their cosets restricting each level in the multi-D mapping. The rotational properties are achieved through the use of a "naturally mapped" 16 QALM signal set. This signal set has the property that, of the four bits used to map the signal set, only two bits are affected by a 90° phase rotation. With an appropriate addition of the coset generators, the multi-D signal sets also have two mapping bits affected by a 90° phase rotation (the remaining bits being unaffected.

This implies that many good rate $k/(k+1)$ trellis codes can be found for effective rates between 3.0 and 3.75 bit/T and that 90° or 180° transparent. The results from a systematic code search using these signal sets are presented. (This work was supported by NASA grant NAG5-557 and by NSF grant NCR89-03429.)

The Extended-*DES*: a Trellis-based Attack Strategy

Jorge M. N. Pereira *Centro de Análise e Processamento de Sinais, Departamento de Engenharia Electrotécnica & Computadores, Instituto Superior Técnico, 1096 Lisboa Codex, Portugal*

In previous work a new and powerful ciphering method based on the *DES* was suggested adding, through **2 Control Bit Streams**, a further security level (Extended-*DES*–E *DES*). Later it was recognized that, given the Key, it would be very easy to recover the *Plaintext* in spite of the additional security. A simple protection strategy was devised consisting of **pre-shifts** of the *Plaintext* and *Ciphertext* (**EE** *DES*). We now describe the attack Strategy (using a trellis approach), and show the security provided by the protected strategy referred to above. Let it be noticed that the additional protection can be controlled by means of the **Control Bit Streams** generation (Pseudo-Noise Bit Sequences) and of the slowly varying pre-shifting **Control Words**. Considering that both the **E** *DES* and the **EE** *DES* can be implemented with a minimum investment and complexity upon the existing *DES* components, the new ciphering system is simple to use remaining extremely versatile.

Performance of Trellis Coded Run-Length Codes

Mignon Belongie and Chris Heegard *School of Electrical Engineering, Cornell University, Phillips Hall, Ithaca, NY 14853*

This paper concerns the performance a family of codes which combines the error correcting capabilities of convolutional codes with run-length constraints. These codes have application to magnetic and optical recording systems.

First, trellis constrained, run-length codes (TCRLCs) are described. The combination of a coset code, C, and (d, k) parameters induce a combined constraint on allowed signals. A signal is a codeword

of the TCRLC if and only if: (1) the sequence of cosets determined by its transition times, t_n, is the coset code, C, and (2) the run-lengths satisfy the given (d, k) run-length constraint, $(d + 1)\Delta \leq T_n \leq (k + 1)\Delta$ where the n^{th} run-length $T_n = t_n - t_{n-1}$ and Δ is the clock period.

After TCRLCs are described, the problem of decoding will be addressed. The decoding problem involves a three step procedure: (1) Detection, (2) Sequence Estimation and (3) "Un-encode". The second procedure, in the case of a recording channel, must make an estimate in the presence of extensive channel memory. This step, which is the most difficult to develop, must trade sequence estimation performance for complexity.

A study is made of the performance of TCRLCs with the described detection methods. In particular, an example the performance of a code that dominates the industry popular "2, 7" and "1, 7" codes (which we term "The Sevens Killer Code") is presented.

Low-Complexity Maximum-Likelihood Decoding Algorithm for Non-Binary Trellis Codes

Gareguin S. Markarian and Haik H. Manukian *Radiophysics & Electronics Institute, Armenian Academy of Sciences, 378410, Ashtarak-2, Armenia, USSR*

In this paper we propose the new simple and fast maximum likelihood decoding algorithm, adapted for non-binary balanced (dc-free) trellis codes. These codes represent a class of codes which map one binary bit into one q-ary ($q > 2$) channel symbol and are used in digital transmission systems, operating by cable line. In order to explain the proposed algorithm we apply the most widely used ternary ($q = 3$) bipolar or the "alternate mark inversion" (AMI) balanced trellis code. The positive effect is achieved by expanding the channel alphabet q to infinity and introducing interdependencies between the channel symbols such that not all vectors of length n with components from expanding alphabet are suitable.

The structure of these codes allows to construct simple and fast soft Viterbi decoder without analog-to-digital converters. In order to interpret the proposed algorithm we use the most widely used ternary trellis "bipolar" or "alternate mark inversion" (AMI) code as an example. The proposed algorithm allows to correct some errors in codes without error correcting properties. The new upper bound for probability of error per symbol is also proposed in this paper. We have found that according to proposed algorithm a theoretical energy gain of 3.5 dB as compared with symbol-by-symbol hard decision detection in AWGN channel is possible. Finally, we present the realization of low-complexity soft Viterbi decoder which proves the operation ability of proposed algorithm.

Performance Evaluation of Trellis Coded Modulations with Memory

Witold Holubowicz and Fidel Morales-Moreno *Technical University of Poznan, 60-965 Poznan, Poland, and Telesat Canada, Gloucester, Ontario, K1B 5P43 Canada*

This paper presents some tools useful in performance evaluation of coded modulations with memory. In the first part of this paper, an efficient method is presented for finding the best combinations of coded modulations with memory, in the sense of maximization of the minimum Euclidean distance d_{min}, without the necessity of doing an exhaustive search. If we want to maximize d_{min} for a class of coded modulations, the search must in general go over all modulations and all codes of interest. In our procedure, used for the optimization of coded Correlative FM signals, we first calculate the upper and lower bounds on the d_{min} for any input sequence with given weight m, for the modulator of Correlative FM without any code. then, we separate all convolutional codes of interest into subclasses, each subclass described by its minimum Hamming distance values. finally, using the previously calculated d_{min} bounds, our search may be limited to one or two subclasses of codes only, still with the same results as those of an exhaustive search.

The second part of our paper deals with ways of calculating upper bounds of the BER for coded modulations with memory. Calculation of the upper bound for a coded memoryless modulation requires, in general, inversion of matrices of size $N2 \times N2$ where N is the number of encoder states. This becomes impractical even for coded modulations without memory. Zehavi and Wolf proposed a method to calculate this bound using only matrices of size $N \times N$, the method however was applicable only to coded

modulations characterized by: 1) linear encoder; 2) code rate $R = (n-1)/n$; 3) memoryless modulation; 4) coded modulations with, so called, "uniform error profiles".

In our paper, we generalize the method described in Zehavi and Wolf by relaxing, step by step, restrictions 2), 3) and 4). In the first two cases, the resulting number of states are still of size $N \times N$. When the error profiles are not uniform, the resulting matrix is larger than $N \times N$, but it is still much smaller than $N^2 \times N^2$. We show simple examples of coded TFM and coded Duobinary MSK, with uniform and non-uniform error profiles, which illustrate the procedures presented earlier.

Multi-Level Multidimensional Trellis Codes

Jiantian Wu and Xuelong Zhu *Department of Electronic Engineering, Tsinghua University, Beijing, P.R. Of China*

Recently the multi-level codes which allow the use of suboptimal multi-stage decoding procedures have well been studied. In this paper, we describe a design method for multi-level codes based on multidimensional signal constellations. We show how to decrease decoding complexity and error coefficients (multiplicity). We present a number of multi-level trellis codes based on four- and eight-dimensional two-way lattice partitions, which have much higher fundamental coding gains than previously known coset codes with moderate decoding complexity. The simulation results of a simple code show that the new code has about the same performance/complexity tradeoff as the previously known best coset codes.

PLENARY SESSION

Thursday, 8 - 8:50 a.m.

Trying to Beat the Heisenberg Principle,

Alberto Grünbaum, *Department of Mathematics, University of California, Berkeley, CA 94720*

TECHNICAL SESSIONS

Thursday, 9 a.m. - 12 m.

SESSION ThA1

ESTIMATION II

Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inverse Problems

Imre Csiszár *Mathematical Institute of the Hungarian Academy of Sciences, Budapest, P.O.B. 127, H-1364, Hungary* (40 min.)

For the set S of all real-valued or all probability mass functions with a given finite domain, all conceivable rules for selecting an element of a feasible subset of S, determined by linear constraints, are considered, and those satisfying certain natural postulates are characterized. Two basic postulates imply that the selection should minimize some function defined on S which, if a prior guess is available, is a measure of distance from the latter. It is shown how invariance properties and a transitivity postulate restrict the class of permissible distances, leading to characterizations of some well-known families of distances and also some new ones. As corollaries, unique characterizations of the methods of least squares and minimum discrimination information are arrived at. The latter are uniquely characterized also by a postulate of composition consistence. As a special case, a unique characterization of the method of maximum entropy from a small set of natural axioms is obtained.

A Method of Sieves for Regularizing Maximum-Likelihood Spectrum Estimates

P. Moulin, D. L. Snyder, and J. A. O'Sullivan *Electronic Systems and Signals Research Laboratory, Department of Electrical Engineering, Washington University, St. Louis, MO 63130*

Maximum-likelihood (ML) spectrum estimation is a notorious ill-posed problem. In this paper, we are concerned with the use of a new regularization method for addressing this fundamental issue. We recommend a method of sieves, based upon the following concepts. The spectrum belongs to a subset of some Hilbert space of functions over which a complete set of nonorthogonal basis functions is defined. The spectrum is then represented by a countable set of coefficients in a nonorthogonal series expansion. By defining an appropriate sieve on this countable set, our problem reduces to maximum-likelihood estimation of the parameters in the sieve. Three main attractive features of this approach are: (1) the nonorthogonal expansion is a convenient framework for defining the sieve and including *a priori* information; (2) mean-square consistence of the estimates can be expected; and (3) we have derived a tractable alternating maximization algorithm for estimating the parameters. The setup of this problem is very general and can be applied without major difficulties to the estimation of higher-dimensional spectral functions.

The Index of Resolvability of Probability Density Estimators

Andrew R. Barron *Departments of Statistics and Electrical & Computer Eng., University of Illinois, 725 S. Wright Street, Champaign, IL 61820*

An index of resolvability is defined which bounds the rate of convergence of density estimators based on the minimum description length (MDL) principle due to Rissanen, Wallace, Sorkin, and others. Given the sample size n, a countable collection Γ_n of probability densities, and codelengths $L_n(q)$, $q \in \Gamma_n$ which satisfy Kraft's inequality, the estimator $\hat{p}_n$ is defined as the minimizer of the total description length $L_n(q) + \log 1/q(X^n)$, $q \in \Gamma_n$. The *index of resolvability* of a density p relative to Γ_n and L_n is defined by

$$R_n(p) = \min_{q \in \Gamma_n} \left\{ \frac{L_n(q)}{n} + D(p\|q) \right\}$$

where $D(p\|q)$ is the relative entropy. This index is seen to bound the redundancy of a universal noiseless source code based on the estimator. Also, it bounds the rate of convergence in squared Hellinger distance, i.e.,

$$d^2(p, \hat{p}_n) \le O_{pr}(R_n(p)).$$

Bounds on the resolvability are determined in parametric and nonparametric cases, yielding near optimum rates of convergence. For instance, estimators based on sequences of approximating exponential families, with the order selected by the MDL criterion, are shown to converge at rate $O_{pr}(n^{-2r/(2r+1)} \log n)$ for log-densities with r square-integrable derivatives and $O_{pr}(\log n)/n$ for densities in one of the families, without prior knowledge of whether the density is finite or infinite dimensional.

On Estimation of Discrete Hammerstein Systems by the Fourier and Hermite Series Estimates

Adam Krzyzak *Department of Computer Science, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Canada H3G 1M8*

We study the estimation of a single-input, single-output (SISO) discrete Hammerstein system. Such a system consists of a nonlinear, memoryless subsystem followed by a dynamic, linear subsystem. We identify the parameters of the dynamic, linear subsystem. We identify the parameters of the dynamic, linear subsystem by the standard correlation and Newton-Gauss method. The main results concern the estimation of the nonlinear, memoryless subsystem, recovering the nonlinearity using the Fourier and Hermite series regression estimates. We prove the density-free pointwise convergence of the estimates, that is the estimates converge for all input densities. The rates of pointwise convergence are obtained for smooth input densities and for nonlinearities of Lipschitz type. Global convergence and its rate are also studied for a large class of nonlinearities and input densities.

On Estimation of Hammerstein Systems by the Recursive Kernel Regression Estimate

Adam Krzyzak *Department of Computer Science, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Canada H3G 1M8*

In this paper we study the estimation of multi-input, single-output discrete Hammerstein system. Such a system contains a nonlinear, memoryless subsystem followed by a dynamic, linear subsystem. We obtain the impulse response of the dynamic, linear subsystem by the correlation method. We as well estimate coefficients of the ARMA model describing the linear subsystem. The main results concern the estimation of the nonlinear, memoryless subsystem. We impose no conditions on the functional form of the nonlinear subsystem, recovering the nonlinearity using the recursive kernel regression estimate. We prove the distribution-free pointwise and global convergence of the estimate, that is, no conditions are imposed on the input distribution and convergence is proven for virtually all nonlinearities. The rates of pointwise as well as global convergence are obtained for all input distributions and for nonlinearities of Lipschitz type. We also discuss possible applications of the studied nonlinear systems in detection and adaptive control.

A New Lower Bound of Cramér-Rao Type for Quantum State Estimation

Hiroshi Nagaoka *Faculty of Engineering, Hokkaido University, Sapporo, Japan*

Let H be a Hilbert space, and let $\mathsf{S} = \{S_\theta; \theta \in \Theta \subset \mathbb{R}^n\}$ be a family of quantum states on H smoothly parameterized by $\theta = (\theta^1, \ldots, \theta^n)$. We consider the parameter estimation problem for S. A quantity $CR \in \mathbb{R}$ depending on θ is called a *lower bound of Cramér Rao type* if the covariance matrix $V \in \mathbb{R}^{n \times n}$ of any unbiased estimator at the state S_θ always satisfies $\mathrm{Tr}\, GV \ge CR$, where $G \in \mathbb{R}^{n \times n}$ is given positive-definite weight matrix. Well-known bounds are CR_S based on *symmetric logarithmic*

derivatives and CR_R based on *right logarithmic derivatives*. A. S. Holevo constructed another bound CR_H such that $CR_H \geq \max\{CR_S, Cr_R\}$. We introduce a new bound CR_{NEW} for the case $n = 2$, which is based on a fundamental inequality on simultaneous measurements of two noncommutative observables, and show that $CR_{NEW} \geq CR_H$. CR_{NEW} is defined via some minimization that cannot be solved explicitly in general. In the simplest noncommutative case $\dim \mathsf{H} = 2$, however, it is explicitly written as $CR_{NEW} = CR_S + (\det G/\det J)^{1/2}\,\mathrm{Tr}\,\mathrm{Abs}\,S_\theta[L_1, L_2]$, where $\{L_1, L_2\}$ are the symmetric logarithmic derivatives of S at θ and $J = [J_{ij}]$ is a 2×2 matrix such that $J_{ij} = \mathrm{Re}\,\mathrm{Tr}\,S_\theta L_i L_j$. Moreover CR_{NEW} is shown to be the maximum among all the lower bounds of Cramér-Rao type in this case.

Asymptotic and Geometric Procedures for Estimating Correlation and Ambiguity Functions

Edward L. Titlebaum and Sanjay K. Mehta *Department of Electrical Engineering, University of Rochester, Rochester, NY 14627*

In the first section of this paper the method of stationary phase is used to derive an expression for the asymptotic approximation for the crosscorrelation function of two FM signals. The instantaneous frequency curves of these FM signals intersect at only one point in the time-frequency space. In the next section, for signal which are slow varying in amplitude, we derive a simple geometric interpretation for the asymptotic results derived in section 2. The intuitively pleasing result is that the crosscorrelation function between the two FM signals is shown asymptotically to be the square root of an area measured in time-frequency space. This result allows for quick estimation of Correlation and Ambiguity functions.

Statistical Performances of Several Eigen-Structure DOA Estimation Methods

Luo Jingqing and Bao Zheng *Electronic Engineering Institute, Xidain University, Xi'an, P. R. China*

This paper presents some asymptotic statistical performances of several eigen-structure DOA estimation methods, named MUSIC method, Minimum-Norm method, Minimum eigenvector method, and Johnson eigenvector method, from the biases and variances of the parameter estimations. The concept of unbiased signal subspace estimation is introduced. The conclusion is that when the noises are zero mean Gaussian white noises and independent of the signals, unbiased signal subspace estimation is obtained, and all four methods presented above give an unbiased parameter estimation, but their variances are different. The MUSIC method and Johnson eigenvector method have the smallest variances, and the Minimum eigenvector method has the largest variance. Some of these results conflict with the results of other authors. Calculations and graphs are presented to show these properties.

SESSION ThA2

MULTIPLE ACCESS III

A Lower Bound to the Packet Waiting Times in the Infinite Population Multiaccess Channel

Mart L. Molle *Computer Systems Research Institute, University of Toronto, Canada M5S 1A4*

Just as recent analyses of multiaccess conflict resolution algorithms have progressed from (merely) finding their maximum achievable throughput to finding complete delay-throughput response time curves, we now extend the work on algorithm-independent upper bounds to capacity by giving a lower bound to the mean delay-throughput curve. Using a "helpful genie" approach, we reduce the problem from a *distributed* multiaccess system (where individual packet arrival times are not visible to the algorithm) to a *centralized* single server queueing system where bulk service (i.e., group testing) is permitted. These customer arrival times consist of the "real" packet arrival times, augmented with "dummy" points from an independent Poisson process at rate $\lambda(1/p-1)$. Thus, customer arrivals form a Poisson process at rate λ/p, and each customer represents a "real" packet according to an i.i.d. Bernoulli trial with probability p. For p sufficiently large, the (bulk) service discipline that minimizes the mean waiting time in the genie-aided queueing system is easily found using known results for optimal group testing algorithms under the Bernoulli arrival sequence model.

Lower Bound for Packet Delay in Random Multiple Access System

B. S. Tsybakov and N. B. Likhanov *Institute for Problems of Information Transmission, Academy of Sciences of the USSR, 19 Ermolova str. GSP-4, Moscow 101447, USSR*

We consider a packet communications network using a channel with slotted random multiple access (RMA) and ternary feedback $\{\theta_t = 0$ empty, $\theta_t = 1$ success, $\theta_t = 2$ conflict $\}$ about event in slot $[t, t+1)$. It is assumed that there is an infinite number of users and the overall flow of requests for packet transmission is Poisson one with intensity λ.

Here the RMA algorithm is an initial set $B_0 C[0, \infty)$ and a function $B_t = B_t(\theta_0, \ldots, \theta_{t-1})$ defined over all $\theta_i \in \{0, 1, 2\}$, $0 \le i \le t-1$, $t = 0, 1, \cdots$ and taking values of sets B_t from time interval $[0, \infty)$. A ready-for-transmission packet is transmitted in slot $[t, t+1)$ if and only if its moment of generation x is in B_t. The packet delay $d = d(\lambda)$ is defined for a given algorithm as the sum of the mean delays of packets generated in slot $[t, t+1)$ and divided by λ with $t \to \infty$.

The stated problem is to find a lower bound $\mathbf{d}(\lambda)$ [for $d(\lambda)$] that is true for all algorithms. The problem is more general than the upper-bound problem for capacity of an RMA system, since the value of minimum intensity $\lambda = \lambda_0$ (such that $\mathbf{d}(\lambda_0) = \infty$) gives us the upper bound $\bar{C} = \lambda_0$.

We find a lower bound $\mathbf{d}(\lambda)$ and compare it with the mean packet delay for the part-and-try algorithm.

Theory of Packet Reservation Multiple Access

David J. Goodman, Sanjiv Nanda, and Uzi Timor *Wireless Information Networks Laboratory, ECE Department, Rutgers University, Box 909, Piscataway, NJ 08855-0909, and Ministry of Defense, Haifa, Israel*

Packet Reservation Multiple Access (PRMA) makes it possible for wireless voice terminals to share a short range radio channel. As a statistical multiplexer, PRMA is subject to fluctuating packet rates from the ensemble of terminals sharing the channel. Because voice packets require prompt delivery, PRMA responds to congestion by dropping packets delayed beyond a specified time limit.

In this paper, we model the ensemble of terminals as a set of coupled queues and use equilibrium point analysis to evaluate system behavior. Based on the equilibrium values of system state variables, we

derive the probability of packet dropping as a function of the system variables, including design parameters and operating conditions. We also establish conditions for system stability and efficiency.

Numerical calculations based on the theory show close agreement with published results based on computer simulations. They also provide valuable guides to system design.

An Architecture for Very High Speed Packet Switching Systems

R. L. Cruz *Dept. of Electrical & Computer Engineering, University of California at San Diego, La Jolla, CA 92093*

We consider a class of architectures for very high speed packet switches that is based on the concept of error correcting codes.

The inputs and outputs of a switch are labeled by binary strings of length n. The outputs are partitioned into output trunk groups. The function of the switch is to route packets from the inputs to the correct output trunk *group*. Each output trunk group is assigned a binary string of length n, called a codeword. The switch attempts to route a packet to the output whose label is the codeword assigned to the desired output trunk group. Misdirection may occur due to contention; this corresponds to "errors" occurring in the "transmitted codeword". By appropriately partitioning the outputs into output trunk groups, we find that a very reliable and fault tolerant mechanism for routing packets may result.

We consider a specific switch design based on a bufferless omega network. For a simple model for packet arrivals, we find that the optimal partitioning does not correspond to minimum distance decoding with the Hamming metric; in fact the optimal partitioning depends on the utilization of the switch.

Stochastic Monotonicity Properties of Multi-Server Queues with Impatient Customers

Partha P. Bhattacharya and Anthony Ephremides *Electrical Engineering Department & Systems Research Center, University of Maryland, College Park, MD 20742*

We consider multi-server queues in which a customer is lost whenever its waiting time is larger than its (possibly random) deadline. For such systems, important performance measures are the number of (successful) departures and the number of lost customers over a time interval. We establish sufficient conditions on the arrival, service, and deadline processes and on the number of servers, for strong stochastic order monotonicity of these measures as functions of some of the system parameters.

The IFFO Protocols Revisited: An Extension for Integrated Communications

Jeffrey E. Wieselthier and Anthony Ephremides *Information Technology Division, Naval Research Laboratory, Washington, DC 20375, and Electrical Engineering Department, University of Maryland, College Park, MD 20742*

A class of multiple-access protocols for data traffic, known as the Interleaved-Frame Flush Out (IFFO) protocols, were first introduced in 1980. These protocols were shown to be well suited for satellite communication environments. In this paper we show how the IFFO protocols can be modified to accommodate a mixture of voice and data traffic by incorporating a "movable-boundary" mechanism. We describe the analytical model of the protocols in detail. We present exact and approximate methods of analysis that are based on the determination of the equilibrium behavior of infinite Markov chains, and which lead to accurate performance evaluation, or in some cases to bounds on performance. Questions of optimization of bandwidth allocation to the two classes of traffic are also addressed.

Conflict Resolution Algorithms for High Error-Rate Multi-Access Channels

George C. Polyzos and Mart L. Molle *Dept. of Computer Science & Engineering, University of California, San Diego, La Jolla, CA 92093-0114, and Computer Systems Research Institute University of Toronto, Toronto, Canada M5S 1A4*

Previous authors have investigated the performance of random-access Tree Conflict Resolution Algorithms for synchronous broadcast communications channels under a simple model for channel errors. This model, proposed by Massey and also by Vvedenskaya and Tsybakov, deals only with forward errors, and it assumes that stations remain synchronized and receive identical feedback. Furthermore, only the following two types of errors are considered: (i) a single transmission that gets destroyed by noise and is interpreted as a conflict, and (ii) an otherwise idle channel that, because of channel noise, tricks the receivers into believing that a conflict occurred.

It has been reported that, assuming independent errors with probabilities ε and δ, respectively, Tree Algorithms with gated or windowed channel access have capacity zero for $\delta > 0.5$. However, we now show that with a simple modification, i.e., by merely repeating slots that have been reported as conflicts for up to n times (or until an idle or a success is reported), we can achieve non-zero capacities for δ arbitrarily close to 1 by choosing $n \geq -1/\log_2(\delta)$. Furthermore, we provide throughput-delay statistics for these protocols.

The Communication Complexity of Solving a Polynomial Equation

Zhi-Quan Luo and John N. Tsitsiklis *Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge MA 02139*

We consider the problem of evaluating a function (x,y) $(x \in \mathbb{R}^m, y \in \mathbb{R}^n)$ using two processors P_1 and P_2, assuming that processor P_1 (respectively, P_2) has access to input x (respectively, y) and the functional form of f. The processors perform this computation by exchanging messages under the restriction that a message sent by any processor is a continuously differentiable function of the data (x or y) possessed by that processor and the messages that it has already received. We establish a new general lower bound on the communication complexity (i.e., the minimum number of real-valued messages that have to be exchanged, as well as some general properties of optimal protocols. We then apply our results to the case where $f(x,y)$ is defined as a root z of a polynomial equation

$$\sum_{i=0}^{n-1} (x_i + y_i) z^i = 0$$

and obtain a lower bound of n (which matches the obvious upper bound). This is in contrast to the $\Omega(1)$ lower bound obtained by applying earlier results of Abelson. Our results while very intuitive are surprisingly difficult to prove and involve primarily techniques from multivariable calculus.

SESSION ThA3

CRYPTOGRAPHY I

On the Quadratic Spans of De Bruijn Sequences

Agnes Hui Chan and Richard A. Games *The Mitre Corporation, Burlington Road, Bedford, MA 01730*
(40 min.)

The quadratic span of a periodic binary sequence is defined to be the length of the shortest quadratic feedback shift register (FSR) that generates it. This notion generalizes the usual notion of the linear span, which is used to analyze the complexity of pseudorandom sequences. An algorithm for computing the quadratic span of a binary sequence is described. The required increase in quadratic span is determined for the special case of when a discrepancy occurs in a linear FSR that generates an initial portion of a sequence. The quadratic spans of binary De Bruijn sequences are investigated. It is shown that the quadratic span of a De Bruijn sequence of span n is bounded above by $2^n - \frac{1}{2}n(n-1) - 1$ and this bound is attained by the class of a De Bruijn sequences obtained from m-sequences. It is easy to see that a lower bound is $n+1$, but a lower bound of $n+2$ is conjectured. The distributions of quadratic spans of a De Bruijn sequences of span 3, 4, 5, and 6 are presented. It appears that the vast majority of de Bruijn sequences have quadratic spans close to the lower bound.

Cascade Ciphers: The Importance of Being First

Ueli M. Maurer and James L. Massey *Institute for Signal and Information Processing, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland*

We consider the security of cascade ciphers where the keys of the component ciphers are independent. It is shown by a counterexample that the intuitive result, formally stated and proved by Even and Goldreich, that the cascade is at least as strong as the strongest component cipher, is only valid under the uninterestingly restrictive assumption that the enemy cannot exploit information about the plaintext statistics. It is proved that, for quite arbitrary notions of what breaking a cipher means and for any reasonable definition of problem difficulty, the cascade is at least as difficult to break as the first cipher. An obvious consequence of this result is that if the ciphers commute, then the cascade is at least as difficult to break as the most-difficult-to-break component cipher, i.e., the cryptographic chain is at least as strong as its strongest link. It is noted that additive stream ciphers do commute, and this fact is used to suggest a strategy for designing secure practical ciphers.

The Hardness of Solving, with Preprocessing, Two Problems Related to Cryptography

Antoine Lobstein *Centre National de la Recherche Scientifique, URA 820, Telecom Paris, Département Informatique, 46, rue Barrault 75634, Paris 13 (France)*

The two problems **Linear Decoding** and **Subset Sum**, which have given birth to the McEliece and the knapsack public-key cryptosystems, are known to be NP-complete. For the first problem, it has been proved that, even if one knows the linear code in advance and can preprocess it, the existence of a polynomial-time decoding algorithm would imply that the polynomial-time hierarchy collapses at an early stage. We give a new, straightforward proof of this result, and prove that the same holds for **Subset Sum**: even if the knapsack is known in advance and can be preprocessed, there is no polynomial-time algorithm solving it, unless the polynomial hierarchy collapses.

A CDMA Security Scheme Using Bit Inversion

D. Despen and N. K. Huang *Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455*

This paper investigates a code division multiple access (CDMA) communication system security scheme, and evaluates the performance of two sequence types based on this scheme. The sequences to be investigated are the *m*-sequences and the No sequences. The security scheme to be studied is bit inversion, the intentional inverting of bits to confuse an unintended user. The two sequence types will be compared on the basis of effective security and the effect of bit inversion on processing gain (PG).

Sequence Complexity and the Directed Acyclic Word Graph

Cees J. A. Jansen and Dick E. Boekee *Philips USFA B.V., P.O. Box 218, 5600 MD Eindhoven, The Netherlands, and Technical University of Delft, P.O. Box 5031, 2600 GA Delft, The Netherlands*

Blumer's algorithm can be used to build a Directed Acyclic Work Graph (DAWG) in linear time and memory from a given sequence of characters. In this paper we show that Blumer's algorithm can be used very effectively to determine the maximum order (or nonlinearity) complexity profile, as introduced by Jansen, of a given sequence in linear time and memory, where each complexity value is determined sequentially, i.e., after each new character of the sequence. We also show that this algorithm can be used to determine the period of a periodic sequence in linear time and memory. Moreover it appears that the DAWG is an even more efficient means of generating the sequence, given a number of characters, than e.g., the non-linear feedback shift register equivalent of that sequence, as it always needs the least amount of characters to generate the remainder of the sequence.

Run Permuted Sequences

Cees J. A. Jansen and Dick E. Boekee *Philips USFA B.V., P.O. Box 218, 5600 MD Eindhoven, The Netherlands, and Technical University of Delft, P.O. Box 5031, 2600 GA Delft, The Netherlands*

This paper describes the construction of classes of binary sequences, which are obtained by permuting the runs of zeroes and ones of some given periodic binary sequence $\mathbf{s} = (s_0, s_1, \ldots, s_{p-1})^{\infty}$, $s_i \in GF(2)$. Based on the run-length notation of a periodic sequence a large class of sequences is constructed by permuting the runs of zeroes and ones of a DeBruijn sequence of given order. The properties of the sequences in this class, in particular the number of sequences in a class and their complexities, are discussed. It is shown that in this way all DeBruijn sequences of given order are obtained, but also many more sequences with higher complexities, all satisfying Golomb's first and second randomness postulates. Hence, it is demonstrated that Golomb's statement that the number of sequences in this class "is slightly larger" than the number of DeBruijn sequences of order n, is not very careful. It is also shown how to generate the sequences in this class with the use of enumerative coding techniques. The binary sequence generator obtained in this way can be useful for cryptographic purposes, e.g., in streamcipher systems.

An Attack on the Clock Controlled Generator Sequences

Chuan-kun Wu *Department of Applied Mathematics, Xidian University, Xian, P. R. of China*

This paper investigates the complexity and statistical characterization of clock controlled generator sequences, shows its advantage and weakness, and gives an algebraic method to attack this model.

The Linear Recurring Sequences over the Residue Class Ring $Z/(m)$

Li Xiangang *Dept. of Applied Mathematics, Zhengzhou Engineering & Technical Institute, P.O. Box 1001-100, Zhengzhou, Henan, P.R. of China*

We shall deal mainly with the problems of linear recurring sequences over the residue class ring $Z/(m)$, considering first the properties of periods of this class of sequences and proving that in many cases the periods can be determined by the periods of their characteristic polynomials. Then we thoroughly investigate the periods of the polynomials in $Z/(m)[x]$ and show the existence, the construction methods, and the numbers of the polynomials with given periods. In particular, using these results, we conclude that there exist linear recurring sequences of order n over $Z/(p^e)$ with maximum periods, which we call *quasi-m- sequences*, and their number is

$$p^{(2n-1)(e-1)}(1-p^{-n})\frac{\phi(p^n-1)}{n},$$

where p is a prime number, $e \geq 1$ and $\phi(\cdot)$ is Euler's function. Finally, the number $N_f(Z/(p^e))$ of nonshift equivalent sequences generated by the same characteristic polynomial $f(x) \in Z/(p^e)[x]$ and the distribution function $Z_p e(b:\tilde{u})$ of the number of occurrences of $b \in Z/(p^e)$ in a full period of the linear recurring sequence $\tilde{u} = (u_t)$ are discussed. It is proved that if $f(x) \pmod p$ is irreducible and the period of $f(x) \pmod{p^e}$ is $p^r \pi_p(f)$, $(0 \leq r \leq e-1)$, then

$$N_f(Z/(p^e)) = \frac{p^{(e-r)n}-1}{\pi_p(f)} + \left[\frac{p^{en-r}-p^{(e-r)n}}{\pi_p(f)}\right]\left[1-p^{-n}\right](1-p^{-(n-1)})^{-1}+1.$$

It is further proved that for any $b \in Z/(p^e)$

$$\left| Z_{p^e}(b:\tilde{u}) - \frac{S_e}{p^e} \right| \leq (1-\frac{1}{p^e})(\frac{S_e}{S_1})(\frac{S_1}{R_e})^{1/2} p^{en/2},$$

where $\pi_p(f)$ is the period of $f(x) \pmod p$ in $Z/(p)[x]$, R_e is the period of $f(x)$ in $Z/(p^e)[x]$, S_e is the period of $\tilde{u}$ over $Z/(p^e)$, and S_1 is the period of $\tilde{u} \pmod p$ over $Z/(p)$. By using the algorithm given in this paper, it is not difficult to determine $\pi_p(f)$, R_e, S_e, and S_1. Thus from these results the linear recurring sequences over $Z/(m)$ are found to have many of the same beautiful properties as those over finite fields.

SESSION ThA4

SPEECH PROCESSING

Fractional Rate Multi-Tree Speech Coding

Jerry D. Gibson and Wen Whei Chang *Dept. of Electrical Engineering, Texas A&M University, College Station, Texas 77843*

We present both forward and backward adaptive speech coders that operate at 9.6, 12, and 16 kbits/s using integer and fractional rate trees, unweighted and weighted squared error distortion measures, the (M, L) tree search algorithm, and incremental path map symbol release. We introduce the concept of multi-tree source codes and illustrate their advantage over classical, multiple symbol per branch, fractional rate trees for speech coding with deterministic code generators. With a frequency weighted distortion measure, the forward adaptive multi-tree coder produces near toll quality speech at 16 kbits/s, while the backward adaptive 9.6 kbits/s multi-tree coder substantially outperforms adaptive predictive coding and has an encoding delay less than 2 msec. Performance results are presented in terms of unweighted and weighted signal-to-noise ratio and segmental signal-to-noise ratio, sound spectrograms and subjective listening tests.

Smoothed DPCM Codes

Wen Whei Chang and Jerry D. Gibson *Dept. of Electrical Engineering, Texas A&M University, College Station, TX 77843*

Rate distortion theory promises that autoregressive sources can be encoded optimally at small distortions (high rates) by a source coder with infinite encoding delay and zero delay at the decoder. However, for instrumentable systems with finite encoding delay and an unmatched code generator or for operation at low rates, decoding delay may provide a performance increment. The alphabet constrained approach to data compression allows delay at both the encoder and the decoder, and Sethia and Anderson incorporate delay in a tree coder code generator by combining a weighted linear interpolation scheme with DPCM. This system, called interpolative DPCM (IDPCM), was shown to outperform DPCM at rate 1 for several synthetic source models. In the present work, we use minimum mean squared error (MMSE) fixed-lag smoothing in conjunction with DPCM to develop a code generator employing delayed decoding. This smoothed DPCM (SDPCM) code generator is compared to DPCM and IDPCM code generators at rates 1 and 2 for tree coding several synthetic sources and to a DPCM code generator at rate 2 for speech sources. The (M, L) algorithm is used for tree searching, and SDPCM outperforms IDPCM and DPCM at rate 2 for the synthetic sources with $M = 1, 4, 8$, and 12, and at rate 1 with $M \geq 4$. For speech, SDPCM provides a slight improvement in MSE over DPCM codes that is also evident in sound spectrograms and subjective listening tests. The models upon which the fixed-lag smoother is based must be chosen appropriately to achieve good SDPCM performance.

Deconvolution of Voiced Speech Based on Minimum-Phase/All-Pass Decomposition

Ki Yong Lee, Iickho Song, and Souguil Ann *Department of Electronics Engineering, Seoul National University, Seoul 151-742, Korea, and Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Seoul 130-650, Korea*

In this paper we view a speech signal as the convolved signal of the impulse response of a non-minimum phase vocal tract and a multi-pulse excitation source. A non-minimum phase can be decomposed into a minimum-phase and an all-pass component. Applying the linear predictive filter for the minimum-phase component and the spiking filter for the all-pass component, we propose a method for deconvolution of the multi-pulse excitation source and the vocal tract. The vocal tract is in the cascade form of a linear predictive filter and a spiking filter. Through computer simulation, we show the

efficiency of the proposed method using the convolved signal of a known non-minimum phase system and a multi-pulse excitation source and using real speech.

Modeling of Speech Excitation Source by a Bernoulli-Gaussian Process

Ki Yong Lee, Iickho Song, and Souguil Ann *Dept. of Electronics Engineering, Seoul National University, Seoul 151-742; Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology, Seoul 130-650; and Dept. of Electronics Engineering, Seoul National University, Seoul 151-742, Korea*

In speech signal an excitation source consists of a few distinct sparse impulse train. In this paper the excitation source is modeled statistically as a zero mean Bernoulli-Gaussian process. The pulse locations are independently distributed with a probability distribution and pulse amplitudes are Gaussian random variables with zero mean and finite variance. In this paper we propose an algorithm using a likelihood function for estimation of pulse amplitudes and locations based on Bernoulli-Gaussian process. To obtain pulse location first we use the cross-correlation function between the speech signal and the vocal tract impulse response. Exploiting these pulse locations, and using a procedure in which the amplitudes of the pulses are kept optimum at each stage, we estimate optimum pulse amplitudes. Through computer simulation, we show that the proposed Bernoulli-Gaussian process model of excitation source is also a resonable model for the deconvolution procedure of real speech if we use maximum-likelihood estimate.

On the Use of Mean and Difference of Adjacent Line Spectrum Pair Frequencies for Speaker Recognition

Chi-Shi Liu, Min-Tau Lin, Wern-Jyuhn Wang, and Jung-Juey Chen *Basic Research Department, Telecommunication Laboratories, Ministry of Communication, P.O. Box 71, Chung-Li Taiwan, 320 R.O.C.*

The line spectrum pair (LSP) was first introduced by Itakura as an alternative linear prediction (LP) representation in frequency domain. In the last years, a number of studies have shown that the LSP representation has more efficient encoding than other LP parametric representations and also provide better recognition rate than other spectral information in speech recognition. But still no papers discuss the role of LSP frequencies in speaker recognition and what performances they have. In this paper, we study the performance of LSP frequencies, and the mean and difference of adjacent LSP frequencies for speaker recognition. The result shows that LSP frequencies, the mean and difference of adjacent LSP frequencies could achieve 98%, 100%, and 100%, respectively by a codebook with 16 codewords and test length with 7 digits.

Predictive Coding for Stationary Gaussian Processes

Kailash Birmiwal *Department of Electrical Engineering, Southern Illinois University, Carbondale, IL 62901*

We consider predictive coding for stationary Gaussian processes and present a rate-distortion formulation of the coding problem. A simple novel iterative algorithm is developed for designing the entropy-encoded quantizer of the predictive encoder for all rates. The algorithm assumes Gaussian distribution as the input to the quantizer, and the parameters of the distribution are updated at each stage. The performance of the algorithm is evaluated numerically, and the results are compared with the optimum distortion-rate functions.

A Space-Variant Covariance Model for DC/AC-Separated Image Block Coding

Yonggang Du *Institut für Elektrische Nachrichtentechnik, Aachen University of Technology, D-5100 Aachen, West Germany*

For blockwise image source coding each picture is partitioned into rectangular blocks of equal size. The statistics, especially the covariance matrix of the blocks, have widely been assumed to be stationary within the block field. In this paper we will show that this assumption is only conditionally correct. It is true if the blocks are processed along with their DC-value. In many practical applications, however, the mean value (DC-value) of each block is often removed from the block and handled in its own manner, because of its weak correlations to the AC-values. One has in fact a DC/AC-separated block coding. In this case the actual blocks to be modeled are those whose DC-term is removed. For such mean-removed blocks it is now shown that their covariance matrix is no longer space-invariant. A mathematical proof is presented in this contribution, confirmed by experimental measures. Model functions are derived for variable block size and for blocks with or without classification in orientation angles.

Gain Adapted Hidden Markov Models for Recognition of Clean and Noisy Speech

Yariv Ephraim *Speech Research Department, AT&T Bell Laboratories, Murray Hill, NJ 07974*

A key issue in applying hidden Markov modeling for recognition of speech signals is the matching of the energy contour of the signal to the energy contour assumed by the model. A mismatch between the energy contours could cause a major problem for speech recognition, since the probability of speech events could be miscalculated by the model. When clean signals are available for recognition, the gain matching is usually achieved by appropriately normalizing the signal prior to both training and recognition. When only noisy speech signals are available for recognition, the standard approach of gain normalization is not applicable. In this paper, a unified approach is developed for designing models for gain normalized signals, and for gain adaptation in recognition of clean and noisy speech signals. During training, the gain function of the speech signal is estimated along with the parameter set of the model. During recognition, joint gain estimation and word decoding is performed. The gain function and parameter set are exclusively estimated by the maximum likelihood estimation approach using the EM (estimation-maximization) algorithm.

SESSION ThA5

BLOCK CODING

Constructions of Error-Correcting DC-Block Codes

Tuvi Etzion *Computer Science Dept., Technion, Haifa 32000, Israel*

A $(2n,l,c,d)$ dc-free binary block code is a code of length $2n$, constant weight n, maximum runlength of a symbol l, maximum accumulated charge c, and minimum distance d. The requirements are that l and c be small. We present two dc-free codes with distance $2d$, $d \geq 1$, length $2m+2r(d-1)$ for $d \leq 3$ and length $2m+2r(d-1)(2d-1)$ for $d > 3$, where $r \leq [\log_2(2m+1)]$. For the first code $l=4$, $c=2$, and the asymptotic rate of this code is 0.7925. For the second code $l=6$, $c=3$, and the asymptotic rate of this code is 0.8858. Asymptotically these rate achieve the channel capacity. For small values of n these codes do not achieve the best rate. As an example for codes of short length with a "good" rate, we first present a (30, 10, 6, 4) dc-free block code with 2^{21} codewords. Finally, we present a construction for which from a given code C_1 of length n, even weight, and distance 4, we obtain a $(4n, l, c, 4)$ dc-free block code C_2, where l is 4, 5, or 6, and c is not greater than $n+1$, but usually significantly smaller. The codes obtained by this method have good rates for small lengths. We discuss the encoding and decoding procedures for all the codes.

A Class of Error and Erasure Control (d,k) Block Codes

H. C. Ferreira and Shu Lin *Laboratory for Cybernetics, Rand Afrikaans University, P.O. Box 524, 2000 Johannesburg, South Africa, and Department of Electrical Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822*

We show that a unique integer composition can be associated with each (d,k) sequence. By imposing some restrictions on such a composition, block codes capable of error and erasure control can be synthesized. We first present three families of restricted (d,k) sequences capable of single error detection. If we append a small and fixed number of parity and buffer bits to these restricted (d,k) sequences of arbitrary length, codes capable of correcting single errors and up to $(d+1)$ adjacent erasures, can be constructed. Codes capable of correcting all single and double adjacent errors and codes capable of correcting up to $(2d+2)$ adjacent erasures, can be constructed either by imposing further compositional restrictions or by increasing the number of parity bits. For each composition, we present difference equations for enumeration, characteristic polynomials and tables with numerical capacities. The rates of the class of codes investigated here, approach the capacity of the d constrained channel for large d. In comparison to recording codes currently used, some of our constructions exceed the minimum separation between transitions in the channel waveform, at the expense of a smaller detection window width.

On Error-Controlling (d,k) Constrained Block Codes

Øyvind Ytrehus *Dept. of Informatics, Thormøhlengst. 55, N-5006 Bergen, Norway*

Binary (d,k) constrained, possibly dc-free, error detecting or correcting block codes are considered with regard to three types of channel errors: binary-symmetric errors, asymmetric errors and bitshift errors. Bounds on the code sizes are obtained from computer search techniques, constructions and sphere-packing bounds. Short block codes can be used as building blocks in longer concatenated (and generalized concatenated) code constructions.

On Multi-level Block Modulation Codes

Tadao Kasami, Toyoo Takata, Toru Fujiwara, and Shu Lin *Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan, and Department of Electrical Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822*

One of the dramatic developments in bandwidth-efficient communications over the past few years is the introduction and rapid application of combined coding and bandwidth-efficient modulation, known as coded modulation. Using coded modulation, reliable data transmission can be achieved without compromising bandwidth efficiency. In this paper, we investigate the powerful multi-level technique for combining block coding and modulation. This multi-level technique allows us to construct bandwidth-efficient block modulation codes with arbitrary large minimum squared Euclidean distances from Hamming distance component codes (binary or nonbinary) in conjunction with proper signal mapping. The paper consists of four parts. In the first part, we present a formulation for signal sets on which modulation codes are to be constructed. Distance measures on a signal set are defined and their properties are developed. In the second part, we present a general formulation for multi-level modulation codes in terms of component codes with appropriate Euclidean distances. The distance properties, Euclidean weight distribution and linear structure of multi-level modulation codes are investigated. In the third part, several specific methods for constructing multi-level modulation codes are proposed. Based on these methods, some short block codes for 8-PSK, 16-PSK and 16-QASK modulations are constructed. These codes have good minimum squared Euclidean distances and provide significant coding gains over some uncoded reference modulation systems with little or no bandwidth expansion. In the last part, error performance of block modulation codes is analyzed for an AWGN channel based on a soft-decision maximum likelihood decoding. Error probabilities of some specific codes are evaluated based on their Euclidean weight distributions.

On Linear Structure and Phase Rotation Invariant Properties of Block 2^l-ary PSK Modulation Codes

Tadao Kasami, Toyoo Takata, Toru Fujiwara, and Shu Lin *Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan, and Department of Electrical Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822*

As the application of coded modulation in bandwidth-efficient communications grows, there is a need of better understanding of the structural properties of modulation codes, especially those properties which are useful in: error performance analysis, implementation of optimum (or suboptimum) decoders, efficient resolution of carrier-phase ambiguity, and construction of better codes. In this paper, we investigate two important structural properties of block 2^l-ary PSK modulation codes, namely: linear structure and phase symmetry. For an AWGN channel, the error performance of a modulation code depends on its squared Euclidean distance distribution. Linear structure of a code makes the error performance analysis much easier. Phase symmetry of a code is important in resolving carrier-phase ambiguity and ensuring rapid carrier-phase resynchronization after temporary loss of synchronization. It is desirable for a modulation code to have as many phase symmetries as possible. In this paper, we first represent a 2^l-ary modulation code as a code with symbols from the integer group, $G = \{0, 1, ..., 2^l - 1\}$, under the modulo-$2^l$ addition. Then we define the linear structure of a block 2^l-ary PSK modulation code over G with respect to the modulo-2^l vector addition, and derive conditions under which a block 2^l-ary PSK modulation code is linear. Once the linear structure is developed, we study phase symmetry of a block 2^l-ary PSK modulation code. In particular, we derive a necessary and sufficient condition for a block 2^l-ary PSK modulation code, which is linear as a binary code, to be invariant under $180°/2^{l-h}$ phase rotation, for $1 \leq h \leq l$. Finally, a list of short 8-PSK and 16-PSK modulation codes is given together with their linear structure and the smallest phase rotation for which a code is invariant.

A Class of Block Codes with Redundant Signal-Sets for PSK-Modulation

Magnus Isaksson and Lars H. Zetterberg *Telecommunication Theory, Royal Institute of Technology, S-100 44 Stockholm, Sweden*

This paper deals with channel codes where the redundancy is obtained not from parity symbols, but from expanding the channel signal-set. They were initially proposed by Ungerboeck using a convolutional code. In this paper, we give a block coding approach.

The expanded signal-set is given the structure of a finite field. The code is defined by a square nonsingular circulant generator matrix over the field. Binary data is mapped on a dataword, of the same length as the codewords, over an additive sub-group of the field. A computer search has resulted in, e.g., rate 2/3 coded 8-PSK of lengths 4-8 with coding gains 0.7-3.0 dB over uncoded QPSK.

We describe the codes using *trellises*, and then apply the *Viterbi algorithm* for decoding. For 8-PSK, there are codes with gain 3 dB with only 4 states in the trellis.

We also show that parallel transitions in the trellis can be described by an additive group, specified by the inverse generator matrix, and its cosets.

More on the Behavior of Binary Block Codes at Low Signal-to-Noise Ratios

Chi-chao Chao and Robert J. McEliece *Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30043, R.O.C., and Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125*

It is well-known that for block codes of a given rate, the larger the minimum distance, the better the code will perform at high signal-to-noise ratios. For some applications such as deep space communications, we care about the behavior of codes at low signal-to-noise ratios. In a former paper by Swanson, McEliece, and Chao, an expressions was derived for the block error probability of binary block codes on the unquantized additive white Gaussian noise (AWGN) channel near the point where the signal-to-noise ratio is zero. In this paper we find a similar expression for the bit error probability with maximum-likelihood decoding. Examples of codes such as orthogonal codes, bi-orthogonal codes, the (24,12) extended Golary code, and the (15,6) expurgated BCH code are discussed. The asymptotic coding gain on the unquantized AWGN channel at low signal-to-noise ratios is also studied. (This work was supported by the Air Force Office of Scientific Research under grant # AFOSR-88-0247.)

Block Coded Modulation on AWGN and Fading Channels

Lin Zhang and Branka Vucetic *School of Electrical Engineering, University of Sydney, NSW 2006, Australia*

In this paper, new BCM codes for 8-PSK modulation with good error rate performance are presented. The performance of these codes on the additive white Gaussian channel (AWGN) and the Rayleigh fading channel is evaluated.

A fading channel, with its specific properties, requires different criteria in constructing short BCM codes from those used for an AWGN channel. Although minimum squared Euclidean distance (d_{free}) is still very important, it plays a less significant role when dealing with severe facing channels. Instead, effective length (l_{eff}) along the path is considered for code construction. Our results show that the 8-state BCM code without parallel transitions in its trellis diagram is superior to those with parallel transitions.

SESSION ThA6

CODING THEORY VI

A Truncated-Stack Sequential Decoding Algorithm: Analysis and Implementation

Pierre Lavoie, David Haccoun, and Yvon Savaria *Department of Electrical Engineering, Ecole Polytechnique de Montréal, P.O. Box 6079, station "A", Montréal (Quebec) Canada*

A new sequential decoding algorithm called "Truncated-Stack Algorithm" (TSA) is presented and shown to be attractive for high speed decoding of long convolutional codes. In this algorithm, a significant reduction in stack size is achieved by gradually discarding from the stack some paths that are very unlikely to be correct. Computer simulations show that the cumulative distribution of the computational effort is the same for the TSA with stack size S and the usual stack algorithm as long as the number of computations remains smaller than a critical value $C_{crit}(S)$. An analysis based on the theory of branching random processes confirms that $S \ll C_{crit}(S)$, implying that the TSA can be more economical to implement than the usual stack algorithm.

The problem of implementing TSA decoders for high-speed applications is examined. A multiprocessor VLSI architecture based on a new systolic priority queue is presented. The new systolic priority queue allows a very fast ordering of the nodes in the stack, hence making TSA attractive for applications requiring powerful error correction at very high data rates.

Sequential Decoding Without a Cut-off Rate

J. B. Anderson, *ECSE Department, Rensselaer Poly. Inst., Troy, NY 12180-3590*

Traditional sequential decoders are designed from the premise that correct decoding of a data symbol must occur with certainty. A standard analysis then calculates the expected number of code tree branches that are visited in the decoding of each symbol. Some difficulties arise in this kind of design. The branch expectation is infinite at code rates above R_0, the computational cut-off rate. The decoder will violate any hard limit to storage or computing power with at least some probability, even at rates below R_0; it then fails to decode correctly, in conflict with the premise. We investigate decoders that are designed from the premise that they may fail to decode correctly with probability $P_e > 0$. Suppose that all code tree paths are traced on a plot of distance to the received path vs. Path length. It turns out that a P_e-decoder must explore all the paths between two boundaries, a drop line and a stop line. Paths may be dropped that hit the drop line, and any path hitting the stop line stops the search; its initial symbol is the decoder output. We give simple expressions for these two lines and for the expected branches visited. These are given for the optimal code searches of two types, those that allow backtracking in the search and those that do not. The channel is the binary symmetric channel with crossover p. Neither type of search shows a cut-off rate phenomenon; the expected branches visited simply grows to infinity as p tends to p_c, the crossover probability leading to a capacity equal to R. The expectation is of the form $\exp[[1+(\log\sigma)/h]\log P_e]$ for searches that backtrack and $\exp[[(\log\sigma)/h]\log P_e]$ for those that do not. Here h and σ are functions of p and R. Furthermore, P_e and the maximum depth of search n are related by $P_e = \exp[-n\,E(R)]$, where $E(R)$ is the Gallager reliability exponent for block codes. The behavior of these P_e-decoders can be directly related to existing decoders such as the stack, Fano, M-, and Viterbi algorithms.

Sequential Decoding and Wald's Identity

Rolf Johannesson and Kamil Sh. Zigangirov *Department of Information Theory, University of Lund, Box 118, S-221 00 Lund, Sweden, and Institute for Problems of Information Transmission, USSR Academy of Sciences, 19 Ermolovoy st., Moscow GSP-4, USSR 101447*

In this paper we give a unified analysis of sequential decoding. All results are based on Wald's identity. First we derive an upper bound on the average number of computations per branch for the stack algorithm. A similar bound is shown for the Fano algorithm. For finite random tree codes terminated by tails of dummy information zeros we give an upper bound on the error probability which is valid for both algorithms. For infinite trees we introduce a fixed backsearch limit and derive an upper bound on the error probability which holds for the ensemble of fixed convolutional codes if the memory of the encoder $m = n - 1$, where n is the backsearch limit. Finally, the different choices of the bias term in the metric are discussed.

Orphans of the First Order Reed-Muller Codes

Richard A. Brualdi and Vera S. Pless *Dept. of Mathematics, University of Wisconsin, Madison, Wisconsin 53706, and Dept. of Mathematics, University of Illinois at Chicago, Chicago, Illinois, 60680*

If C is a code, an orphan is a coset which is not a descendant. Orphans arise naturally in the investigation of the covering radius. In case C has only even weight vectors and minimum distance at least 4, we characterize cosets which are orphans, and then prove the existence of a family of orphans of first-order Reed-Muller codes $R(1,m)$. For $m \leq 5$ all orphans of $R(1,m)$ are identified.

Coset Codes with Isometric Labelings

G. David Forney, Jr. *Codex Corporation, 20 Cabot Boulevard, Mansfield, MA 02048*

A coset code $\mathbb{C}(\Lambda/\Lambda'; C)$ based on a lattice partition Λ/Λ' results from mapping the sequences $\mathbf{y}$ in some group code C to sequences of cosets of Λ' via a labeling function that maps code outputs (labels) y_k to cosets of Λ'. If the labeling function is linear, then $\mathbb{C}$ is linear. Isometric labelings are a broader class that lead to codes ('pseudolinear codes') that have geometrical invariance properties similar to those of linear codes. In particular, a pseudolinear code is distance-invariant, and all of its Voronoi regions have the same shape. These properties extend to 'label translates' of pseudolinear codes. All known good lattice-type coset codes are pseudolinear. Linear labelings, or translates of linear labelings, are isometric, and isometric labelings are regular. A binary isometric labeling exists if and only if $\mathbf{Z}^N/\Lambda/\Lambda'/4\mathbf{Z}^N$ is a lattice partition chain, whereas a linear labeling exists if and only if $\mathbf{Z}^N/\Lambda/\Lambda'/2\mathbf{Z}^N$ is a lattice partition chain (for some scaling of Λ/Λ'). Furthermore, if there is an isometric labeling for Λ/Λ', then there is an isometric Ungerboeck labeling for Λ/Λ'. If $\mathbb{C}$ is a trellis code based on a linear time-invariant convolutional code C, then the time-zero lattice of $\mathbb{C}$ is a pseudolinear code Λ_0, and the set of possible outputs from any encoder state is a label translate of Λ_0.

Operating Cosets in Arbitrary Lattice Partitions

Mauro A. O. da Costa e Silva, *Faculdade de Engenharia Eletrica, Universidade Estadual de Campinas, Caixa Postal 6101, Campinas-SP, Brasil*

In the computational search for optimum trellis coded modulation (TCM) schemes based on lattice partitioning, as proposed by Calderbank and Sloane, it is convenient to generate an addition table for a set of coset representatives, each time a different partition Γ/Λ of a lattice Γ by a sublattice $\Lambda \subseteq \Gamma$ is used, in order to simplify the search for good generalized convolutional codes based on Γ/Λ, whose output coset is a linear combination of some cosets in Γ/Λ chosen as the columns of its generator matrix. A special selection of coset representatives for arbitrary lattice partitions is proposed and a closed analytical expression for addition of these coset representatives is derived. Besides its application in the search for

optimum TCM schemes, the material could be used to verify whether a given periodic packing is indeed a lattice.

New Lower Bounds for Asymmetric Codes

Tuvi Etzion *Computer Science Department, Technion - Israel Inst. of Technology, Haifa 32000, Israel*

We present three different methods to obtain lower bounds for binary asymmetric codes. The first method is by using some combinations of the Cartesian product between a partition of all the binary n-tuples into disjoint constant weight codes with distance 4 and a partition of all the binary k-tuples into disjoint asymmetric codes with distance 2. By using this method we obtain lower bounds for binary error-correcting codes of length $n+k$ with asymmetric distance 2. The second method is by combining the Preparata code of length 2^n (or the shortened Preparata code of length 2^n-1), the extended Hamming code (or the Hamming code), and some binary k-tuples. By using this method we obtain codes of length 2^n+k (or 2^n+k-1) and asymmetric distance 3. the third method is by considering the weight distribution of known codes and their translates. In a number of cases the results significantly improve on the best lower bounds previously known.

Sequential Decoding with an Incremental Redundancy ARQ Scheme

S. Kallel *Department of Electrical Engineering, University of British Columbia, 2356 Main Mall, Vancouver, B.C., Canada V6T 1W5*

In this paper, sequential decoding is analyzed in conjunction with an efficient incremental redundancy ARQ scheme using punctured convolutional coding. With the incremental redundancy ARQ scheme, whenever the decoding time for a given data packet exceeds some predetermined value T_{max}, decoding of that packet is stopped and incremental redundancy bits are provided by the transmitter, decreasing thus the coding rate. Should decoding still fail, then the transmitter sends additional incremental redundancy bits, decreasing once again the coding rate. This procedure continues until decoding finally succeeds.

It is shown that with sequential decoding used in conjunction with the incremental redundancy ARQ scheme, the throughput increases as the starting coding rate increases. Moreover, it is shown that with a starting high coding rate code, the throughput is always better than with a rate 1/2 code with code combining. Thus, the incremental redundancy ARQ scheme makes sequential decoding very powerful, allowing the communication system to be flexible and adaptive to channel conditions, even under wide noise variations and severe channel degradations.

SESSION ThA7

VITERBI DECODERS

Windows, Multipath and Viterbi Modems

Torleiv Maseng and Odd Trandem *Elab-Runit, N-7034 Trondheim, Norway*

A Viterbi decoder is used to demodulate a signal corrupted by noise and multipath. A new technique is presented for choosing the part of the channel impulse response which minimizes the intersymbol interference caused by a lengthy channel response, considering the finite complexity of the receiver. This enables longer multipath profiles to be equalized without increasing the number of states. It is efficient when applied to nonlinear modulation techniques with compact power spectra. Such nonlinear modulation schemes are of interest because they may contain inherent coding gain. These schemes are exemplified.

Error Computation of Viterbi Decoder

H. F. Rashvand *Technophone Ltd., Research Dept., Surrey, GU15 3SP, England*

A new model for calculating the output error probability of a convolutional code that uses the Viterbi Decoding Algorithm (VAD) is introduced. Its structure is based upon the Markov chains model of the system. It makes use of a state transition matrix and an error generating vector. Then from this model two programmable algorithms are derived corresponding to two different implementations of the VAD. In the first case the error generating vector is initially set at the error free condition and then taken through the trellis while adjusted step-by-step by applying a transition matrix. The vector follows a chain of identical successive processes that eventually reaches a final state for the specified search length. The error computation of the code then becomes a straightforward accumulation of an error contributing subset of the vector. For the second case, however, first a state-metric transition matrix is generated. This matrix is then converted into an error generating vector that is adjusted according to decoder's search length. The rest of this algorithm is similar to the first case. The accuracy of the algorithms is compared with results of the Monte Carlo Method and illustrated against some upper and lower bounds.

A Burst Error Model of Viterbi Decoding for BPSK Modulation on Fading and Scintillating Channels

Joel M. Morris and Deval Patel *Electrical Engineering Department, University of Maryland Baltimore County, Baltimore, MD 21228*

The burst error statistics of a soft-decision Viterbi decoder were simulated and a corresponding model obtained for the case when the transmitted signal is: (1) encoded with the 313 (3, 1/2), 31123 (5, 1/2), or 3233013 (7, 1/2) convolutional codes; (2) modulated via coherent BPSK for the AWGN channel; and (3) subjected to slow and non-selective scintillation/fading modeled by the Nakagami-*m* distribution. These statistics were generated by Monte-Carlo simulations, and compiled in terms of burst-error-length average and quantile (90%, and 99%) statistics vs SNR (E_b/N_o) parameterized by the fading intensity parameter *m*. These statistics are evaluated to determine the appropriateness of extending the geometric burst-error model for the nonfading case to the case of Nakagami-*m* fading. The results have important implications for the design of interleaved or non-interleaved concatenated coding schemes, and to the development of burst error models and simulators for scintillating and fading channel environments.

Generalized Viterbi Algorithms (GVA) for the Decoding of Convolutional Codes

Nambirajan Seshadri and Carl-Erik W. Sundberg *AT&T Bell Laboratories, Murray Hill, NJ 07974*

Two generalizations of the Viterbi algorithm (VA) are presented for the decoding of convolutional codes. They are respectively, a parallel version that simultaneously identifies the L globally best candidates ($L > 1$), and a serial version that iteratively finds the L best candidates. An erasure is declared if the metric difference of the best and the second best path falls below a certain threshold. Such an information is useful in combined source and channel coding, where an unreliable frame of decoded data can be replaced using inter-frame source redundancy either through the process of prediction or interpolation. Alternatively, the inter-frame source redundancy can be used to select the best out of the L candidates that are released by the channel decoder. In data transmission schemes, an outer encoder adds intra-frame data redundancy to the data to be encoded by the inner convolutional code. At the receiver, the inner decoder, which is based on the GVA, produces a list of the L best candidates, and the outer decoder selects that candidate that satisfies the intra-frame data redundancy check. Gains of more than 1.0 dB are possible using GVA with $L = 3$ over the $VA(L = 1)$ on the AWGN channel. The modulation is BPSK and the inner code is a rate $R = 1/2$, memory $M = 4$ convolutional code. For the Rayleigh fading channel, a gain of more than 2 dB is obtained. Signal space geometry ideas are used to illustrate how these gains are achieved. For the AWGN channel, a lower bound on the asymptotic coding gain for different values of L are presented for the convolutional codes by using the simplex signal set.

Multistage Decoding Using a Soft-Output Viterbi Algorithm

Joachim Hagenauer and Peter Hoeher *German Aerospace Research Establishment (DLR), Institute of Communications Technology, D-8031 Oberpfaffenhofen, West-Germany*

The Viterbi algorithm (VA) is modified to deliver not only the most likely path sequence in a finite-state Markov chain, but either a probability estimate or an analog ('soft') value for each symbol. This reliability indicator allows for making soft decisions used for decoding outer codes. In such a way the inner VA accepts and delivers soft sample values and can be regarded as a device for improving the SNR, similar to an FM demodulator. Several applications of multistage decoding are investigated to show the gain over the conventional hard-deciding VA including concatenated convolutional codes. The Soft-Output Viterbi Algorithm (SOVA) can also be applied to the concatenation of convolutional and simple block codes, the concatenation of trellis-coded modulation and convolutional FEC codes, and to coded Viterbi equalization. For these applications we found additional gains of 1 to 4 dB as compared to the classical hard-deciding algorithms.

Quantization Effects in Viterbi Decoding

Ivan M. Onyszchuk, Kar-Ming Cheung, and Oliver Collins *California Institute of Technology and Jet Propulsion Laboratory, Pasadena, CA 91125*

In order to design efficient Viterbi decoders, the tradeoff between the number of input quantization bits q and performance loss must be known. An optimal (uniform) quantization method for the AWGN channel is presented. The corresponding channel capacity $C_u(q)$ is used to determine the smallest value of q that does not substantially increase a $q = \infty$ decoder's bit error rate (BER). The quantizer stepsize which maximizes $C_u(q)$ almost minimizes the decoder BER. The range of state metrics is analyzed to determine a small state metric register width. We describe a simple renormalization scheme that will be used in JPL's new $K = 15$, rate 1/6, multichip decoder for the *Galileo* deep-space mission. These results yield design parameters for reduced hardware complexity Viterbi decoders with negligible BER increase from maximum likelihood decoding on the AWGN channel.

High Speed Viterbi Decodings Having an Idle Mode

Kazuhiko Yamaguchi and Hideki Imai *Department of Computer Science and Information Mathematics, University of Electro-Communication and Division of Electrical and Computer Engineering, Yokohama National University, Yokohama, Japan*

This paper is concerned with high speed Viterbi decoding algorithm having two modes (i.e., *'active mode'* and *'idle mode'*). By the evaluation of the algorithm, we show a new criterion of good convolutional codes for the algorithm. The basic concept of the algorithm is that the Viterbi decoding is performed only when errors are detected. Even if we use a component of the slow-speed Viterbi decoder, the algorithm can attain high communication speed. For example, the case of rate 1/2 convolutional code with 6 encoder memories realize 100 times higher communication speed than that of the component Viterbi decoder at the point of channel bit error rate 10^{-3}. By using the transfer functions, we prove the algorithm is guaranteed to be maximum-likelihood decoding (MLD) for all noncatastrophic convolutional codes. From the evaluation of the condition to be MLD for each convolutional code, we give the suitable generator polynomials of convolution code for the algorithm. The analysis method is related to the study of the error probability of Viterbi decoding with truncated memory. The interesting result is given in the presentation.

High Speed Viterbi Decoder Structures

Erik Paaske *Institute of Circuit Theory and Telecommunications, 343, Technical University of Denmark, DK-2800 Lyngby, Denmark*

We propose new Viterbi decoder structures which exploits the more powerful unit memory or partial unit memory codes, decodes more than one bit per clock cycle and have only few operations per decoded bit. They are suitable for implementation with advanced standard integrated circuits, in particular large memory chips. Data rates from 150 Mbit/s to more than 300 Mbit/s ware achievable without using multiplexed decoders.

The structure is based on the state space of the syndrome former. The state metrics and decisions in level $l + 1$ can be expressed by

$$\mathbf{M}_{l+1} = f(\mathbf{M}_{ell}, \mathbf{S}_{l+1}) \text{ and } \mathbf{D}_{l+1} = g(\mathbf{M}_{ell}, \mathbf{S}_{l+1})$$

where $\mathbf{M}_l$ denote State metrics at level l and S_{l+1} the syndrome generated by the received bits. By appropriate renormalizing of $\mathbf{M}_l$ the number of metric combinations N_c is final, and if it is small enough, the functions in (1) can be performed by a single ROM which replaces all ACS circuits in a conventional decoder. For hard decisions powerful codes directly suitable for such structures are found and for soft decisions, where N_c are generally quite large, methods are developed to reduce N_c in such way that no or only a small performance degradation results.

TECHNICAL SESSIONS

Thursday, 2 p.m. - 5 p.m.

SESSION ThP1

DETECTION THEORY II

Fast Simulation of Detector Error Probabilities in the Presence of Memory and Non-Linearity

Randall K. Bahr and James A. Bucklew *Department of Electrical & Computer Engineering, University of Arizona, Tucson, AZ 85721, and Department of Electrical & Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706*

One would like to compare and analyze *digital communication systems* based upon their overall probability of error. Unfortunately, useful closed form expressions for these probabilities are almost impossible to derive due to the complexity of the stochastic system. Hence one must resort to simulation techniques. One such technique is *Monte Carlo* simulation, which directly counts the number of errors in repeated trials. The error probabilities are usually quite small, requiring numerous simulation runs to sufficiently "hit" the rare event, a bit or symbol error, to gain adequate knowledge of its statistics. This places severe demands on the computer's random number generator. *Importance sampling* strategies simulate under altered input signal distributions (translation or stretching) so as to "speedup" convergence of the error estimators. In this paper we discuss a speedup technique based upon the observation that the error estimator's variance vanishes exponentially fast as sample size increases. The simulation is carried out under a new input signal distribution which is adjusted to maximize the rate of exponential decrease. The fast simulation method is compared to two other *importance sampling* techniques currently in use.

Distributed Detection with Decision Feedback

Rajan Srinivasan *Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY 13244-1240*

A distributed detection scheme using feedback is proposed and studied. On each observation, peripheral detector decisions are combined in a central processor. Feedback links from center to each peripheral exist. Using a simple traffic-reduction protocol, the central decision is communicated to all peripheral detectors and serves to update peripheral decisions on a new observation. The sequence of hypothesis tests representing the decision process is shown to be consistent. It is argued that overall link usage can be appreciably reduced for this scheme. Performance results for a Rayleigh channel are presented and suggestions for further investigations made.

Optimum Detection in the Presence of Random Transient Disturbance and White Gaussian Noise

T. T. Kadota *AT&T Bell Laboratories, Murray Hill, NJ 07974*

Detection of a deterministic signal in white Gaussian noise and a random transient disturbance is considered. Instead of the usual likelihood ratio as the detection statistic, we propose the use of the conditional likelihood ratio given the disturbance where the disturbance is estimated by the maximum a posteriori likelihood method. Since the likelihood ratio is the conditional version averaged over all possible realizations of the disturbance but the disturbance in the data, if any, is one particular realization, use of the conditional version is more appropriate. The disturbance model chosen is a piece of Gaussian process multiplied by a Rayleigh-distributed constant where the starting time and the duration are uniformly distributed. Such a disturbance, together with white noise, constitutes the worst case of the contaminated Gaussian noise in the min-max robust detection. The detection procedure in the presence of

this disturbance consists of two steps: (i) first to determine whether or not the disturbance is present and, if it is present, (ii) then to estimate it and subtract the estimate from the data before the matched-filtering. These two steps tend to be computation-intensive and various ad hoc but more efficient schemes are being investigated.

The Use of ARE in Finite Sample Size Detector Performance Prediction

Rick Blum and Saleem A. Kassam *General Electric Co., and Department of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104*

The limitations of the asymptotic relative efficiency (ARE) in predicting finite sample size detector performance have been noted in several previous studies. It has been observed that the finite sample size relative efficiency (RE) will often converge very slowly to its asymptotic limit. In addition, in some cases the RE approaches the ARE from below while in other cases the RE overshoots the ARE and approaches it from above. It is demonstrated in this paper that both of these effects can be predicted for a large and useful class of detectors used for detection of known signals in additive noise. The prediction formula is developed through analysis for a general detector structure. The result is then used to analyze some specific detectors. Among the specific detectors are the sign detector, the dead-zone detector, the four-level detector, and the more general N-level quantized detector. In addition, the soft-limiter and the noise blanker are analyzed. Computer calculations and simulations giving the relative efficiency as a function of sample size have been used to verify our predictions. These results are also presented.

Reduced State Sequence Detection of Continuous Phase Modulation

Arne Svensson *Ericsson Radar Electronics AB, Aerospace Division, Airborne Radar Systems Design, S-431 84 Mölndal, Sweden*

A reduced state sequence detector (RSSD) which combines decision feedback or hard decision with Viterbi decoding is proposed for Continuous Phase Modulation. With decision feedback, this detector can successfully be used for both binary and high-level CPM schemes. This detector is analyzed by means of minimum Euclidean distance and by simulations of the symbol error probability.

It is shown that many CPM schemes can be decoded without using phase states in the decoder without sacrificing asymptotic error performance (minimum distance). For some schemes, especially high-level schemes, the number of states can be further reduced while still maintaining asymptotically optimum error performance. Furthermore, a very small number of states can be employed in the decoder, with a relatively small degradation in the error performance. With hard decision, a detector not making use of phase states can be used for small modulation indices $1/P$ with a minor degradation in error performance.

Quickest Detection of Time-Varying Signals

S. D. Blostein *Department of Electrical Engineering, Queen's University, Kingston, Ontario K7L 3N6, Canada*

In the past, research in quickest-detection has been confined to the detection of a change from one stationary stochastic process to another. In many situations, signals of interest may be time-varying, as in the problem in detecting slowly emerging targets in radar clutter or in rapidly detecting sampled periodic signals such as for carrier synchronization in coherent communications systems. A quickest-detection algorithm for time-varying changes in the means of a discretized data sequence is presented in this paper. This is accomplished by generalizing reaction procedures studied by Page and Lorden to the time-varying case. Bounds on false alarm rate are derived using a generalized version of a theorem by Lorden and applied to the test design problem of threshold selection. In addition, an upper bound of the average reaction time of the quickest detection algorithm is derived for the case of white Gaussian noise. To test the theory, the analytical results are compared to simulations performed on discretized step-functions, ramps, and periodic functions of known frequency in white Gaussian noise.

Threshold Signal Detection and Estimation in Signal-Dependent Noise

David Middleton *127 East 91 Street, New York, NY 10128*

Recent analyses of threshold signal detection (D) and estimation (E) involving generalized noise and interference have dealt primarily with "telecommunication" environments where the significant noise is *ambient*, i.e., does not depend on the original transmitted signal. However, for many applications additional scattered or (unresolvable) signal-dependent components occur, often accompanied by (resolvable) multipath. This feature, of course, has been extensively considered in earlier work, but under gauss noise conditions and usually without explicit physical structure. But for the many instances of small-angle, high-frequency operation, and where only a few significant scatterers are involved, among other conditions, this noise is no longer normal, and can be highly nongaussian.

Accordingly, it is necessary to extend the earlier canonical theory. The present paper accomplishes this by appropriately treating the scatter component as an additional signal. This also lowers the model's statistical burden, since only second-order second-moments (covariances), at most, of these new "signals", are required. (The first-order pdf of the ambient noise, as before, is still needed in these independent sampling cases.) The resulting threshold algorithms and performance measures are noticeably modified. Performance may be enhanced or degraded, depending on the statistics of the scattering model. This is illustrated by examples, including radar and sonar, as well as telecommunication applications.

Weak Signal Detection in Correlated Non-Gaussian Noise

David Middleton *127 East 91 Street, New York, NY 10128*

Previous work of the author and others in developing an effective canonical theory of weak-signal detection in nongaussian noise has required the assumption of independent sampling. While this is useful even when there is weak correlation in the noise samples, the Locally Optimum Bayes Detectors and Estimators derived from it are not strictly optimum nor is the associated performance. Moreover, although independent time samples are easy to obtain, independent spatial sampling is more difficult, because of the strongly correlated nature of most noise fields.

Efforts to overcome this difficulty have been made by a number of authors recently. A more general analysis, based on "moving-average" noise models, following the results of the author, has been carried out by Maras, using fully probabilistic criteria of performance. However, his analytical results do not lead readily to numerical results for performance and algorithms. The present paper overcomes these limitations, particularly for the common cases of weak correlation in the noise samples $0(\geq 0.2\text{–}0.3)$, by exploiting the usual Töplitz nature of the noise covariances. Added new features are various explicit exact results in important special cases. Both broad- and narrow-band signals and noise are treated. Calculations show typical improvement is 0(2 dB) with sample correlations 0(0.2), or 0(10) in error rates. (Work supported by ITS/NTIA of U.S. Dept. of Commerce, Boulder, CO, 80303.)

SESSION ThP2

COMMUNICATION NETWORKS

Code Combining with Convolutional Coding and Sequential Decoding for CDMA Slotted Networks

T. Ketseoglou and A. Polydoros *Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-0272*

We apply the idea of code combining to CDMA slotted networks with convolutional coding and sequential decoding. Both hard and soft decision decoding are examined. In addition, we study two different decoding scenarios. Under the first, perfect channel state (side) information (number of simultaneous transmissions) is available at the receiver (the decoder operates under perfect channel match) while under the second scenario no such information is available (the decoder operates under channel mismatch). Our results indicate high improvement in the average utilization and delay, at the expense of modest increase in system complexity.

Multi-Access Protocols for Metropolitan Area Networks

Manoel A. Rodrigues *AT&T Bell Laboratories, Holmdel, NJ 07733*

We propose two distributed multi-access protocols and show that they operate well even in the presence of propagation delays. We consider a slotted bi-direction Bus being shared by bursty users. Each user is assumed to know the direction in which to transmit a given message. A limited number of flags can be send in the header of slots flowing in the direction opposite to transmission. In the first protocol, a user sends a busy signal when its state changes to busy (i.e., having something to transmit) and an idle signal when it becomes idle. In the second protocol, a user sends a busy signal when it changes to a busy state and an idle signal M slots later, where M denotes the size of the message to be transmitted. Upon observing an idle signal, a user estimates the number of active downstream users. Slots are allocated to emulate a target service discipline (e.g., round-robin or FIFO). We analyze our two protocols and the IEEE 802.6 MAN protocol, and show that the proposed protocols are more suited to MANs.

Optimal Admission and Routing at a Simple Data Network Node

Ioannis Lambadaris *Electrical Engineering Department, University of Maryland, College Park, MD 20783*

A stream of messages reaches the buffer of a node in a communication network, according to a Poisson distribution with parameter λ. The messages are then to be routed over two channels with different propagation times that are exponentially distributed with parameters μ_1 and μ_2 where μ_1 is larger than μ_2. Two controls are to be applied at the node. The first control concerns the admission of a message into the buffer while the second control is responsible for routing the buffered messages over the two channels. We determine the optimal policy which minimizes an expected cost involving the weighted sum of a penalty for not admitting a message into the buffer and a penalty for the delay experienced by the messages already stored in the buffer. (This research was supported by the Systems Research Center at the University of Maryland under NSF grant CDR-88-03012.)

Performance Evaluation of Multi-Access Strategies for an Integrated Voice/Data Packet Radio Network

Mohsen Soroushnejad and Evaggelos Geraniotis *Department of Electrical Engineering & Systems Research Center, University of Maryland, College, Park, MD 20742*

The problem of voice/data integration in a random-access radio network employing the ALOHA protocol in conjunction with recursive retransmission control is investigated. Code division multiplexing

is used as a suitable modulation in a radio environment to decrease the effect of multiple-access interference. Multi-access control strategies are introduced which take advantage of multiple-access capability of the CDMA channel to accommodate several voice calls simultaneously, while the data users contend for the remaining (if any) multiple-access capability of the CDMA channel. The retransmission probabilities of the backlogged data users are updated based on estimates of data backlog and number of established voice calls which are obtained from the side information about the state of channel activities. A two-dimensional Markovian model is developed for the voice and data traffic, with the data backlog and number of established voice calls representing the state of the system. Based on this model, the voice-call blocking probability, the throughput of both traffic types, and the delay of the data packets are evaluated and the tradeoffs between the parameters of different traffic types are quantified.

Stability Analysis of Asymmetric, Limited Service, Polling Systems

Ramesh Rao and Amir Behroozi-Toosi *University of California, San Diego, La Jolla, CA 92093*

Three cyclic server strategies have been studied in the literature. They are the exhaustive service strategy the gated service strategy and the limited service strategy. In the performance analysis of such systems much of the attention has been given to delay analysis of the system and the system is usually assumed to be stable or in a stationary state. With gated or exhaustive service strategies, if one of the queues is unstable then all other queues become unstable too. This is not the case for the limited service strategy.

In this work we present a method for deriving the stability conditions of limited service systems by combining a result from Lyones and a technique of constructing a sequence of auxiliary systems which dominate the actual system in a well defined sense. The stability region of the auxiliary system is derived without resorting to stationarity assumptions. Thus we are able to derive inner bounds to the stability region of the actual system.

Avoiding Third Order Intermodulation Interference in Mobil Radio Systems

Torleiv Kløve *Department of Informatics, University of Bergen, Thormøhlengst. 55, N-5008 Bergen, Norway*

We consider mobile radio system for a collection of areas, and without third order intermodulation interference within each area. The problem has been considered by a number of authors, and it requires the following construction:

An (I,J)-set of Disjoint Distinct Difference sets (DDD) is a set

$$\Lambda = \{\Lambda_1, \Lambda_2, \ldots, \Lambda_I\}$$

where

$$\Lambda_i = \{a_{ij} \mid 1 \le j \le J\} \text{ for } 1 \le i \le I$$

are disjoint sets of positive integers such that, for each i, all the differences $a_{ij} - a_{ij'}$, with $j' \ne j$ are distinct.

Let

$$h = h(\Lambda) = \max\{a_{ij} \mid 1 \le i \le I,\ 1 \le j \le J\}$$

For the application we want an (I,J) DDD with h as small as possible. Let

$$H(I,J) = \min\{h(\Lambda) \mid \Lambda \text{ is an } (I,J) \text{ DDD}\}$$

Clearly $H(I,J) \ge IJ$. We will present constructions of DDD which shows that $H(I,J) = IJ$ for $I \ge 4J$. The construction is based on a generalization of sonar sequences

Delay and Throughput in a Frequency-Hop Communication Network

Sang Wu Kim *Korea Institute of Technology, 400 Kusung-dong, Yusung-gu, Taejon 305-701, Korea*

The packet delay measured in average number of retransmissions before the successful transmission of a packet in a slotted frequency-hop packet radio network is considered. Reed-Solomon coding is employed to correct the errors or the erasures occurring due to the multiple-access interference. We obtain the packet delay in terms of code rate and channel traffic, and determine the constraint on them for the delay to be less than a specific value. We also derive an accurate approximation to the constraint. The throughput of the network along with the constraint on delay is investigated to see the tradeoffs in the parameters of the network. Both side information and no side information cases are considered.

It is observed that the delay increases abruptly after some channel traffic, but is shown to be kept within a bound by controlling the code rate. This indicates that we can make the network stable by controlling the code rate according to the channel traffic. The delay performance is found to be tremendously improved by using side information at the receiver, and this improvement becomes more significant as the channel traffic increases. The maximum channel throughput is found to be attained at very low values of delay (typically when delay ≤ 1), and sacrificing the delay does not necessarily increase the channel througput.

A New Decoding Scheme for Convolutionally Coded ARQ

T. Hashimoto *Dept. of Information Sciences, Tokyo Denki Univ., Hatoyama, Saitama 350-03, Japan*

The Viterbi algorithm finds out the maximum likelihood path by comparing, at each step and for each state, paths that merge on the given state (or node) in the trellis diagram. The algorithm tends to make an error when the likelihood of the correct path is as small as incorrect ones due to channel noise.

Known ARQ schemes based on the Viterbi algorithm employ devices which request the retransmission until one path, which is supposed to be the correct one, shows a sufficiently greater likelihood than the remainders. Since those ARQ schemes can not detect errors occurred in the past, the error probability depends on how incorrect decisions are prevented at each step. Thus, efforts to attain higher reliability inevitably increase the probability of retransmission.

In this paper, we propose a new ARQ scheme based on the generalized Viterbi algorithm (GVA) and show that the same reliability is attainable with the same complexity, but with less retransmission probability. Actually, we can make the error probability arbitrarily small by increasing the length of the tail appended to each frame without substantially increasing the decoder complexity.

A Sub-Optimal Distributed Self-Organizing Mobile Radio Network Algorithm

Yong Li and Bing-Zheng Xu *South China University of Technology, Guangzhou 510641, P. R. China*

In this paper we have proposed and analyzed a sub-optimal distributed algorithm for self-organization networks, which can organize a set of scattered, mobile, radio-equipped nodes into a connected network. The algorithm has the ability to reorganize the nodes into a new network under very unfavourable circumstances (jamming, wartime conditions). This feature can improve the survivability of the network. It uses network topology information obtained from the partial connectivity matrix that can yield to fewer control nodes and which can reduce the overhead traffic and improve the responsiveness of the network to changing traffic conditions. Using Slotted-ALOHA as channel access scheme, we analyzed and compared the original and new proposed algorithm, and found the throughput of the sub-optimal self-organization network to be better. The computer simulation of the algorithm has also been given.

SESSION ThP3

NEURAL NETWORKS II

Analysis of a Modified Hebbian Rule

Ph. Piret *Philips Research Laboratory, Av. Van Becelaere 2, Box 8, B-1170 Brussels, Belgium*

Let X be an $m \times n$ matrix over $U = \{+1, -1\}$ and associate to X its Hebb matrix $H = X^T X$. Let U_0 be the set $\{+1, 0, -1\}$. It is well known that H can be used to define a mapping $U^n \to U_0^n : y \mapsto \text{sgn}\ (yH)$, where sgn denotes the signature of its argument. In this paper one analyzes the properties of a modified mapping $U^n \to U_0^n : y \mapsto \text{sgn}\ (yG)$, where G is itself the signature sgn (H) of H. Assuming that $(m-1)$ is close to $(n-1)\beta/\log(n-1)$ on determines first for which values of β the mapping $y \mapsto \text{sgn}\ (yG)$ stabilizes the rows of X. In particular, one shows that

- for any $\beta \geq 0$, the (r,s) entry of sgn (XG) is most often equal to the (r,s) entry of X,
- for $\beta < 1/\pi$, the r^{th} row of sgn (XG) is most often equal to the r^{th} row of X,
- for $\beta < 1/2\pi$, the matrix sgn (XG) is most often equal to X.

Then one obtains upper bounds on the (mean) residual error probability when the mapping $y \mapsto \text{sgn}\ (yG)$ is used to map y on the row of X that is closest to y. Although there does not exist any $\tau > 0$ such that for $n \to \infty$ one guarantees that *all* errors in $\lfloor \tau n \rfloor$ positions of a row of X are correctable in one step, one can prove that *most* of them are correctable by the mapping $y \mapsto \text{sgn}\ (yG)$ when τ is smaller than a positive quantity, which is a function of β.

An Iterative Learning Algorithm That Updates Only When It Learns

S. C. Huang and Y. F. Huang *Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556*

The primary objective of this paper is to investigate a learning algorithm, recently proposed by Huang and Huang, for perceptron-like networks. the main feature of this algorithm so-called SUBPA is that it implements a selective learning strategy, thus updates only when it can learn from the input data.

Results of convergence analysis for single layer, as well as multi-layer, networks will be presented. It is shown that, under very general conditions, the algorithm converges after a finite number of updates.

An associative memory machine (AMM) is devised here using perceptron networks. This machine is novel as it is constructed with perceptron networks, as opposed to Hopfield networks. In addition, it can be used to demonstrate the capability of the aforementioned learning algorithm. The notion of *global realization* is introduced here. It is shown that SUBPA is particularly suitable for applications to global realization. The AMM is also used to demonstrate the class of problems for which the back propagation algorithm can be suitable. Furthermore, applications to data compression for image processing purposes using perceptron-like networks are also demonstrated.

Image Prediction for Error Concealment Using Neural Networks

Nobukazu Doi, Toshiaki Takahashi, and Hideki Imai *Central Research Laboratory, Hitachi, Ltd., Tokai Works, Hitachi, Ltd., and Faculty of Engineering, Yokohama National University, Japan*

An image prediction method is proposed for error concealment using a neural network. A (48,9,1) three-layered feed-forward network is chosen. The input layer is connected to pixels around an erroneous pixel to be concealed. The output layer has one unit and puts out a signal that is predicted to be capable of concealing the erroneous pixel.

Using several images as training input and output data, the back-propagation learning procedure adjusts the network's link weights to achieve ideal predictions. Image simulation confirms a neural network is better than the conventional linear prediction method, especially for high-frequency images.

However, there is a possibility of errors occurring in pixels connected to the input layer's units. To cope with such cases, a hybrid neural network is proposed that uses signals obtained by the linear prediction method instead of erroneous pixel signals. It's confirmed that the hybrid method performs better than a linear filter only.

Neural Network Applications for Jamming State Information Generator

Lawrence T. Schaefer and Hyuck M. Kwon *Lockheed Missiles and Space Center Company, Sunnyvale, CA 94088 and Houston, TX 77058, and Electrical Engineering and Computer Science Department, The University of Wisconsin-Milwaukee, WI 53211.*

A previous jamming state information (JSI) scheme, which was proposed for a coded frequency-hopped MFSK (FH/MFSK) system under partial-band noise jamming plus additive white Gaussian noise, utilized the maximum a posteriori (MAP) rule based on the total energy received in the M-tone signaling bands. It was assumed that the knowledge of partial-band noise jamming fraction was available to that JSI generator. Because this scheme reduces the M-dimensional information contained within an M-dimensional signal vector into one dimension, i.e., the total energy, the generated JSI may not be the best. In addition, this scheme is hard to implement because the MAP rule needs the $M-1$ order modified Bessel function. In this paper we present a neural network approach to the JSI generation which is not only implementable but also uses the innate characteristic of the M-dimensional vectored data to create the JSI. The efficiency of the new JSI generator with known partial-band noise jamming fraction will be compared with the MAP generator. Afterwards, the neural network scheme will be generalized to increase its robustness by allowing for an unknown partial-band noise jamming fraction. We find that the neural network JSI generator with or even without knowledge of jamming fraction can improve significantly the performance of a coded FH/MFSK communication system over the MAP JSI generator for high code rate. For example, the neural network JSI generator without knowledge of jamming fraction is 1.43 dB better than the MAP JSI generator in the bit energy-to-jamming noise power spectral density ratio required to achieve the cutoff rate of 0.7, for FH/binary FSK with the symbol energy-to-thermal noise ratio 15 dB.

An Optimal Neural Net Model for Image Coding in The Position-Frequency Space in the Presence of Noise

A. M. Elramsisi and M. A. Zohdy *School of Engineering and Computer Science, Oakland University, Rochester, MI 48309-4401*

Image representation in the frequency-position space (joint domain) is investigated based on the optimal properties of the Gabor basis functions (GBF's). A major intent of this article is to embed the joint domain optimal features of GBF's in the structure of neural networks for image representations in the presence of noise. The voltage controlled oscillator (VCO) neuron model is first introduced. Its joint domain neuron characteristic function (NCF) is analyzed through the investigation of the Kolmogorov-Smoluchowski steady state differential equation (K-S DE) of the model for an optical input signal. At higher values of signal-to-noise ratio (SNR), the NCF function of the VCO neuron model resembles the GBF, and reflects their joint domain optimal properties. The range of resolution and consequently the trade-off between the spatial domain and frequency domain resolutions is now dictated by SNR values. The proposed neuron model is then encompassed in a three-layered neural network for image representation in the joint domain. The scheme is later expanded to account for nonuniform sampling as a characteristic feature of human vision. The performance of the proposed scheme is evaluated in terms of the fidelity in reconstructing the images after extracting the coefficients of expansion. Simulation results, modification to improve convergence of the network, and possible extensions for this work are also provided.

On the Stochastic Approximation Based Learning Algorithm for Neural Networks

Valadimir Nedeljkovic and Milan Milosavljevic *Institute of Applied Mathematics and Electronics, Belgrade, Yugoslavia, and School of Electrical Engineering, University of Belgrade, Belgrade, Yugoslavia*

In this paper we consider the RHW (rumelhart, Hinton, Williams) neural networks in the statistical pattern recognition tasks. We propose the stochastic back-propagation algorithm (SBP) as a modification of the well known error back-propagation algorithm (BP. The central part of our work is the proof of the theorem which gives sufficient conditions for convergence of the SBP algorithm, in mean square and with probability 1, to a local minimum of the criterion function. The final part of this work presents experimental results of comparison of SBP and BP using test examples suggested by Kohonen, Barna, and Chrisley. Comparative analysis of our experimental results with that reported by Kohonen, Barna, and Chrisley confirm superiority of SBP algorithm both in convergence rate as well as in accuracy of recognition.

On Segmentation and Recognition of Connected Digits Based on Neural Network Model

Jhing-Fa Wang, Chung-Hsien Wu, Ruey-Chinq Shyu, and Jau-Yien Lee *Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.*

In this paper, an automatic segmentation and recognition system based on a neural network model is proposed and back-propagation learning algorithm is employed to establish these networks.

The main new idea for segmentation is to classify the energy, spectrum, and pitch-period transitions that occur at the boundaries between syllables. These feature transitions are used as the input patterns of the neural network. The segmented syllables are then used as the basic units for recognition by feeding them into the well-trained network.

Ten digits (0-9) and syllables spoken in Mandarin are used in the speaker-independent phase for segmentation experiments, and only ten digits are used in speaker-dependent phase for recognition experiments. With an average speaking rate at 150 per minute, the segmentation accuracy measured in coincidence rate of 95.7% and recognition rate of 97.2% can be achieved.

Back-Propagation Neural Network Model Based On-Line Chinese Character Recognition System

I-Chang Jou and Ching-Feng Hsu *Telecommunication Labs., Ministry of Communications, P.O. Box 71, Chung-Li, Taiwan, R.O.C., and Institute of Information & Electronics Engineering, National Central University, Chung-Li, Taiwan, R.O.C.*

A neural-network-based on-line Chinese character recognition (OLCCR) system is presented in this paper. The OLCCR system is able to recognize a Chinese character immediately after it is written on a tablet. The recognition process is composed of two phases: (1) the feature matching phase, in which a feature vector based on the stroke information is extracted, and (2) the character matching phase, in which the feature vector is matched with feature vectors of standard Chinese characters stored in the reference pattern database. In this paper, back-propagation neural-nets models, instead of conventional classification algorithms, are used to solve the pattern matching problems in OLCCR. Experimentally, the time to recognize a handwritten character is 0.15 second on the average, and the recognition rate from 73.3% to 94.4%, depending on the user's writing style.

SESSION ThP4

CRYPTOGRAPHY II

A Zero Knowledge Proof Protocol for Communications Authentication

Jonathan D. Low *Verifone, Inc., 100 Kahelu Avenue, Mililani, HI 96789*

This paper describes a procedure for verifying the identity of a remote system. The confidence level of the verification approaches 100% rapidly with time or space. The algorithm uses only counting and parity testing and therefore executes effectively. The intelligence bits transmitted between the two parties have an equal probability of appearing as a 1 or 0 to the enemy. Hence, the transmitted bit sequences actually carry zero information.

Communication Complexity of Secure Distributed Computation in the Presence of Noise

Eytan Modiano and Anthony Ephremides *Department of Electrical Engineering, The University of Maryland, College Park, MD 20742*

We consider a simple model of distributed computation that requires information exchange over a noisy channel. We utilize a communication protocol that requires alternate bit exchanges between two processors and are interested in determining the communication complexity of this exchange. First, we consider the case of a single public channel and compute the number of bits that need to be exchanged between the processors to permit δ-accuracy in their goal. For this computation we consider an error-detection-and-retransmission mechanism of error control as well as an error-correction-and-retransmission mixture that are consistent with the logical protocol that governs this exchange. Second, we consider the case of the availability of an additional secret channel and are interested in determining the minimum number of bits that need to be exchanged over a secret channel in order to maintain ϵ-uncertainty about the computation for an eavesdropper on the public channel. We consider various subcases under this case and obtain an upper bound on the number of secret bits when no error-control scheme is used.

Modified Graham-Shamir Knapsack Public-Key Cryptosystem

Chi-Sung Laih, Lein Harn, Jau-Yien Lee, and Yan-Kuin Su *Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C., and Computer Science and Telecommunications Program, University of Missouri-Kansas City, Kansas City, MO 64110*

Since Shamir first proposed the polynomial time attack to the basic Merkle-Hellman knapsack cryptosystem, lately almost all knapsack cryptosystems based on the modular multiplications can be attacked. This is due to the fact that those systems must satisfy the constraint $M > \sum_{i=1}^{n} a_i$, where a_i's are the deciphering sequences and M is the modulus. In this paper, we propose a modified version of Graham-Shamir knapsack public-key cryptosystem. By appropriately choosing the parameters, one can control the density of the public knapsack, which is the ratio between the number of elements and the maximum value in bits in the knapsack. It is easy to design a knapsack whose density is high enough to foil "low-density" attacks against our system. Moreover, the constraint of the modular multiplications does not require in the modified version. We call this structure as the "partial structure". Therefore, we believe that no attacks can break this system in a reasonable amount of time at the moment.

On the Secure Communications of Group Oriented Societies

Tzonelih Hwang *Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.*

A practical protocol based on the RSA scheme is developed for large group-oriented networks. The information sender does not have to specify who will receive the message because the members are anonymous and the organization policy of the receiving group is unknown. However, he is able to indicate the nature and destination of the message, e.g., how important the message is and in which order or to which department it will be sent. The receiving group can create its own policy on who can read the message according to the indication of the sender and the group's organization.

Algebraic-Code Cryptosystems for Information Privacy, Reliability and Authenticity

Tzonelih Hwang and T.R.N. Rao *Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C., and University of Southwestern Louisiana, Lafayette, LA 70504*

When confidential data are transmitted over communicational channels or stored in storage systems, one of the most important issues of concern is the assurance of their privacy, reliability and authenticity (integrity). Algebraic codes have long been proved to be very powerful in correcting noise errors. In this paper, algebraic-code encoding is combined with block chaining encryptions into one to obtain data secrecy, data reliability and data authenticity simultaneously. These schemes can be implemented very efficiently, thus are very practical to be used in computer systems.

Unbiased Block Substitution

Lothrop Mittenthal *Teledyne Electronics, 649 Lawrence Drive, Newbury Park, CA 91320*

A block substitution is a one-to-one mapping of the n-bit binary numbers onto themselves. Such substitution or transformation can be represented by a permutation. It is shown that certain permutations of the n-bit binary numbers define a block substitution by modulo 2 addition of one permuted set of numbers to another. Such permutations are termed replicative. A subset of these have an additional feature which is that the equations which they define have an additive relationship when viewed as vectors. Such permutations are termed additive. An unbiased block substitution is defined. It is shown that substitutions defined by an additive permutation are unbiased and that any unbiased block substitution can be represented by a replicative permutation. Additive permutations are shown to form groups which retain the same properties. The conditions for existence of these additive permutations are established, some properties of the groups determined, and the number of such groups enumerated and compared with all possible permutations of the n-bit numbers.

Cryptographic Interleaving

Sami Harari *Université de Toulon et du Var, Faculté des Sciences et Techniques, Avenue de l'Université, 83130 La Garde, France*

The Relationship Between MDS Codes and Threshold Schemes

Yang Yi Xian *Dept. of Information Engineering, P.O. Box 145, Beijing University of Posts and Telecommunications, Beijing, 100088, P.R. of China*

The deep relationship between MDS codes and threshold schemes in cryptography is initially discovered in this paper. We have found that by using the generator matrices or the parity-check matrices of MDS codes, we can implement a new class of threshold systems in cryptography.

A New Measure for Stream Cipher Systems to Defend Against Correlation Attack and Linear Approximation Attack

Zhang Muxiang *Department of Applied Mathematics, Xi'dian University, Xi'an, People's Republic of China*

This paper shows that, in analysis and design of stream cipher systems of several linear feedback shift registers with a nonlinear binary function, there exists not only a trade-off between the nonlinear order and the Correlation-Immune order of the nonlinear binary function, but also a trade-off between the Correlation-Immune order and success-rate of linear approximation to the nonlinear binary function. In order to avoid these trade-offs, a new measure for stream cipher systems to defend against Siegenthaler's correlation attack is presented in this paper.

SESSION ThP5

SHANNON THEORY IV

Non-Equiprobable Signaling on the Gaussian Channel

A. R. Calderbank and L. H. Ozarow *Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill NJ 07974* (40 min.)

Many signaling schemes for the Gaussian channel are based on finite-dimensional lattices. The signal constellation consists of all lattice points within a region R, and the shape of this region determines the average signal power. In the limit as $N \to \infty$, the shape gain of the N-sphere over the N-cube approaches $\frac{\pi e}{6} = 1.53$ db. We show that the full asymptotic shape gain can be realized in any fixed dimension by non-equiprobable signaling. We describe shaping schemes that achieve a significant fraction of the available asymptotic shaping gain. The peak to average power ratio of these schemes is superior to that of equiprobable signaling schemes based on Voronoi regions of multidimensional lattices. the new shaping schemes admit a simple staged demodulation procedure.

The Merit Factor of Binary Sequences Related to Difference Sets

Jørn M. Jensen, H. Elbrønd Jensen, and T. H. Høholdt *Mathematical Institute - 303, The Technical University of Denmark, DK-2800 Lyngby, Denmark*

The merit factor of long binary sequences constructed from cyclic difference sets are investigated. It is shown that only sequences arising from the Hadamard difference sets, that is $(v,k,\ ) = (r5-1, 2t-1, t-1)$, can produce long binary sequences with a nonzero merit factor. Following Baumert, Hadamard difference sets can be classified according to the parameter v. 1) $v = 2^m - 1$: Singer- and GMW-difference sets. 2) v a prime: The quadratic residues modulus v. The Hall-sets if $v = 4x^2 + 27$. 3) $v = p(p+2)$: The twin-prime difference sets. The sequences arising from the Singer difference sets are the maximal-length linear shift register sequences (or m - sequences). The sequences constructed from the quadratic residues are the Legendre sequences. It has been shown that the maximal merit factor of m - sequences, Legendre sequences and twin-prime sequences are 3, 6 and 6 respectively.

Lower Bounds to Moments of List Size

Erdal Arikan *Department of Electrical Engineering, Bilkent University, P.O. Box 8, Maltepe, Ankara, 06572, Turkey*

The list-size random variable L for a block code is defined as the number of incorrect messages that appear to a maximum-likelihood decoder to be at least as likely as the true message. Lower bounds to the moments of L are of interest in a number of applications, particularly in lower-bounding the moments of computation in sequential decoding. We prove the following bounds for the list-size L belonging to a code with M codewords and blocklength N:

$$E(L^t) \geq K^t P_{\min}(M,N,K)$$

$$E(L^n) \geq n!(M/K)^n P_{\min}(K,N,n)$$

where $t \geq 0$ is an arbitrary real number, $n \geq 1$ is an arbitrary integer, K is an integer between 1 and M, and $P_{\min}(M,N,K)$ denotes the minimum, over all codes with parameter (M,N), of the probability of list-of-K decoding error. We then prove by applying sphere-packing lower bounds to the first inequality above that, for any $t \geq 0$,

$$E(L^s) \geq \exp[tN(R-R_t) - o(N)]$$

where $o(N)$ is a quantity that goes to zero as N goes to infinity, and

$$R_t = (1/t) \max_{Q} -\ln \sum_j \left[\sum_i Q(i) P(j \mid i)^{1/(1+t)} \right]^{1+t}$$

where the maximum is over all probability distributions on the channel input alphabet, and the sums are over the channel input and output alphabets, respectively. For sequential decoding, this implies that the t'th moment of computation is unbounded at rates above R_t, for all $t \geq 0$, and settles a long-standing open problem.

Nonuniform Sampling Theorems

Michael David Rawn *Dept. of Mathematical Sciences, Manchester College, North Manchester, Indiana 46962*

This paper takes as point of departure the work of Paley and Wiener, and later Levinson, on nonharmonic Fourier series, to derive nonuniform sampling methods. One theorem yields an extension of the work of Linden and Abramson on sampling using derivative samples. Some basic results of spline theory and Riesz bases in Hilbert Space are used in conjunction with Levinson's work on gap and density theorems to obtain this nonuniform sampling method. In another direction, we present a theorem on nonuniform sampling using entire interpolating functions that are Lebesgue square integrable. This sampling theorem demonstrates - among other things - the stability of the Bessel-type sampling expansions first discussed by Kramer, and by Campbell. General results on closure and completeness of sequences of Bessel functions obtained by Boas and Pollard, identities involving the beta function, and a transformation relating Fourier and Hankel transforms to one another are used in deriving this nonuniform sampling method.

On the Converse Theorem in Statistical Hypothesis Testing

Kenji Nakagawa and Fumio Kanaya *NTT Systems Laboratories, 812C, 1-2356, Take, Yokosuka-Shi, Kanagawa-Ken, 238-03 Japan*

The converse theorem of simple statistical hypothesis testing is investigated according to the Neyman-Pearson theorem and the differential geometry of the space of probability distributions on finite atoms. We obtained the result that the condition $r > D(p_1 \| p_0)$ can be divided into two cases, and a lucid explanation is given for Han and Kobayashi's linear function $f_r(\cdot)$. We also mention general upper bounds on the first-kind error probability. These results can be extended to Markov sources.

Let p_t be the +1 geodesic connecting p_0, p_1 and let $p_\infty \triangleq \lim_{t \to \infty} p_t$. We call $D(p_\infty \| p_0)$ the upper bound of the exponent of the first-kind error probability. We obtain the theorem: if $r > D(p_\infty \| p_0)$, the power exponent is $r + D(p_\infty \| p_1) - D(p_\infty \| p_0)$, which equals $f_r(\cdot)$.

Optimization of Signal Sets for Partial-Response Channels

Michael L. Honig, Kenneth Steiglitz, and Stephen Norman *Bellcore, 445 South St., Morristown, NJ 07960; Dept. of Computer Science, Princeton, University, Princeton, NJ 08854; and Information Systems Laboratory, Stanford University, Stanford, CA 94305*

Given a linear, time-invariant, discrete-time channel with impulse response $h[k]$, we consider the problem of constructing N input signals of finite length T that maximize minimum l_2 distance between pairs of outputs. Two constraints on the input signals are considered: a power constraint on each of the N inputs (*hard constraint*), and an average power constraint over the entire set of inputs (*soft constraint*). The hard constraint problem is equivalent to packing N points in an ellipsoid in min $(T, N-1)$ dimensions to maximize the minimum Euclidean distance between pairs of points.

We describe gradient-based numerical techniques that are used to find locally optimal solutions to the preceding signal design problems with both hard and soft constraints. In the case of hard constraints, feasible descent directions are found by solving linear programs. For the soft constraint problem, we maximize a penalty function that approximates the minimum-distance cost function. The penalty function is similar in form to the error criterion used by Foschini, Gitlin, and Weinstein, where two-dimensional signal sets are optimized for a nondispersive channel with additive white Gaussian noise.

Numerical results, consisting of minimum distance vs. input length for different information rates, are given for the soft constraint problem. The channels considered are the identity channel, the $1-D$ channel, and the $1-D^2$ channel. The computer-generated signal constellations are superior to multi-dimensional constructions based on dense lattices when the number of points per dimension is small.

Our numerical results are compared to upper and lower bounds on ε-capacity, which for a given linear time-invariant channel and information rate gives the maximum achievable minimum distance, or coding gain, as the input length tends to infinity. Results indicate that asymptotically, the maximum coding gains for both the $(1-D)/\sqrt{2}$ and $(1-D^2)/\sqrt{2}$ channels are approximately 1 dB less than the maximum coding gain achievable for the identity channel at a rate of 1 bit/sample, and 2 dB less at a rate of 2 bits/sample.

On Randomization in Communication Complexity

King Fai Pang *Systems Research Laboratory, LSI Logic Corporation, 4400 Bohannon Drive, Suite 230, Menlo Park, CA 94025*

Processors **A** and **B** have n-bit binary sequences **x** and **y** respectively. They communicate with a protocol, over a noiseless channel, so that at least one of them computes the value of a function $\mathbf{f(x,y)}$ for all $\mathbf{x} \in \mathbf{X}$ and $\mathbf{y} \in \mathbf{Y}$. We show that when errors are allowed for a small portion of the inputs, randomization only reduces the number of bits required substantially in the worst case. Specifically, we prove an exponential gap between worst case randomized and deterministic complexities. The average case complexities, on the other hand, only differ by at most a constant factor.

SESSION ThP6

CODING THEORY VII

Error Evaluation for Nonbinary BCH Codes by Lagrange Interpolation

C.P.M.J. Baggen *Philips Research Laboratories, P.O. Box 80.000, 5600JA Eindhoven, The Netherlands*

In this paper an alternative is given for Forney's method of error evaluation that is encountered in time-domain decoding of nonbinary BCH codes. This new method avoids the division needed at that stage of decoding. It turns out that a Lagrange polynomial $L(z)$ can be found such that for all errorlocations X_i, the error value Y_i satisfies : $L(X_i^{-1}) = Y_i$. Furthermore, $L(z)$ can be determined using Euclid's algorithm before the roots of the error locator are known. Therefore all divisions can be concentrated in a processor-like structure that also solves the Key-Equation, leading to reduced hardware implementations.

M-Adic Residue Codes

Vanessa Job *Department of Mathematics, Marymount University, 2807 Glebe Road, Arlington, VA 22207-4299*

The m-adic residue codes, defined by Brualdi and Pless, are a generalization of the quadratic residue codes. They exist at prime lengths p over fields $GF(q)$ when $m \mid (p-1)$ and $(q,p) = 1$. The m-adic residue codes are investigated. Included are the proof of the existence of an infinite class of m-adic residue codes for each m, a construction of a subgroup of the automorphism group for each m-adic residue code, a theorem which gives restrictions on the form of the idempotents of the m-adic residue codes, and lower bounds on the minimum odd weights of the odd-like m-adic residue codes. In addition, containment relations between the m-adic residue codes and the quadratic residue codes are demonstrated.

Tables are included which contain idempotents and minimum weights of the binary m-adic residue codes of lengths less than or equal to 127. Many of these codes have the highest possible minimum weights known for codes of their lengths and dimensions.

Legendre Sums and Weights of QR Codes

Tor Helleseth *Department of Informatics, Thormøhlensgt. 55, N-5006 Bergen, Norway*

We will present a connection between Legendre sums and the weights of the codewords in a circulant code related to a quadratic residue (QR) code. Let $F = GF(p)$, p a prime, denote the finite field with p elements. Let $(\frac{x}{p})$ denote the Legendre symbol defined by

$$(\frac{x}{p}) = \begin{cases} 1 & \text{if } x \in Q \\ -1 & \text{if } x \in N, \\ 0 & \text{if } x = 0, \end{cases}$$

where Q is the set of squares and N the set of nonsquares in F. Let $E(x) = \sum_{q \in Q} x^q$, and let C be the linear circulant code of length p generated by $E(x)$ and it cyclic shifts. An arbitrary codeword in this code can be written as

$$c(x) = x^{j_1} E(x) + \cdots + x^{j_r} E(x) \pmod{x^p - 1}$$

where $0 \le j_1 < \cdots < j_r < p$.

Main theorem. Let w denote the weight of the codeword $c(x)$, then

$$w = \tfrac{1}{2}\left[p + (-1)^{r-1}\left[\sum_{t \in F}\left(\frac{f(t)}{p}\right) - \sum_{i=1}^{r}\left(\frac{g_i(j_i)}{p}\right)\right]\right]$$

where $f(t) = \prod_{k=1}^{r}(t - j_k)$ and $g_i(t) = \dfrac{f(t)}{t - j_i}$.

The results are proved using methods for solving a system of equations in finite fields and by using Gaussian sums.

Bandwidth Efficient Concatenated Schemes for Fading Channels

Branka Vucetic *Electrical Engineering Department, The University of Sydney, Sydney, Australia*

Error control coding can greatly improve the performance and extend the range of fading channels. An important feature of concatenated coding schemes is that very low error rates can be achieved with reasonable complexity of decoding. Concatenated codes consisting of short constraint length trellis inner codes with different number of states and Reed-Solomon outer codes are considered to achieve large coding gains with small bandwidth expansion in the presence of frequency-nonselective slow Rayleigh and Rician fading. Both errors-only and errors-erasures decoding algorithms for outer codes are applied. Upper bounds on bit error probability performance in the presence of fading are obtained and compared with simulation results for zero channel memory. The effect of interleaving in eliminating channel memory is investigated. The performance gains that are achieved by the coding scheme relative to the reference uncoded systems are illustrated via some examples. Results are obtained both by analytical methods and computer simulation.

Balanced Binary Pseudorandom Sequences with Low Periodic Correlation

Shinya Matsufuji, Kyoki Imamura, and Sueyoshi Soejima *Department of Electrical Engineering, Saga University, Saga, 840 Japan, and Department of Computer Science and Electronics, Kyushu Institute of Technology, Iizuka, Fukuoka, 820 Japan*

The families of binary {0, 1} pseudorandom (PR) sequences with low periodic correlation properties necessary for a spread-spectrum multi-access communication system have been known such as Gold, Kasami, bent-function and No-Kumar sequences. These sequences, except the bent-function sequences, are not balanced, i.e., the maximum value of the sequence imbalance (the difference between the number of 1's and the number of 0's in one period) is very large.

In this paper a new family of $2^{n/2}$ binary PR sequences of period $2^n - 1$, where n is even, is introduced and shown that every sequence within the family is balanced, i.e., the sequence imbalance is at most 1.

Periodic Correlation Function of the Bent-Function Sequences

Noriyuki Koga, Shinya Matsufuji, Kyoki Imamura, and Sueyoshi Soejima *Dept. of Elec. Eng., Saga University, Saga, 840 Japan, and Dept. of Comp. Sci. and Elec., Kyushu Institute of Technology, Iizuka, Fukuoka, 820 Japan*

Bent-function sequences of period $2^n - 1$, where $n = 0 \pmod 4$, are balanced binary {0, 1} sequences and have a large linear span. However, the explicit formula have been not known for their periodic auto and cross-correlation functions. Especially there have been no study on the number of phaseshifts at which the correlation function takes the same value. This paper gives the explicit formula for the value distribution of periodic correlation function of the $2^{n/2}$ bent-function sequences $\{f_{\mathbf{A}}(\alpha^i) \mid 0 \le i \le 2^n - 2, \mathbf{A} \in V_{n/2}\}$ defined by $f_{\mathbf{A}}(x) = g_{\mathbf{A}}(\mathbf{u}) + tr_1^n(\sigma x)$ and $g_{\mathbf{A}}(\mathbf{u}) = \mathbf{u}_1^t \mathbf{u}_2 + h(\mathbf{u}_2) + \mathbf{A}^t \mathbf{u}$, where α is a primitive element of $GF(2^n)$, V_k the set of all binary column vectors of dimension k, $\mathbf{u}^t = [\mathbf{u}_1^t\ \mathbf{u}_2^t]$, with $\mathbf{u}_1, \mathbf{u}_2 \in V_{n/4}$, a column vector whose i-th element is defined as $tr_1^n(\beta_i x)$ using a basis $\{\beta_1, \beta_2, \ldots, \beta_{n/2}\}$ for

$GF(2^{n/2})$ over $GF(2)$, $tr_1^n(\,)$ the trace from $GF(2^n)$ to $GF(2)$, $h(\,)$ an arbitrary function mapping $V_{n/4}$ into V_1, $\sigma \varepsilon\, GF(2^n)GF(2^{n/2})$, and the superscript t denotes the transposition.

Improved Balanced Encoding

R. M. Capocelli, L. Gargano, G. Landi, and U. Vaccaro *Dipartimento di Matematica G. Castelnuovo, Universita' di Roma "La Sapienza", 00185 Roma, Italy, and Dipartimento di Informatica ed Applicazioni, Universita' di Salerno, 84081 Baronissi (SA), Italy*

A binary code of length n is called balanced if each codeword contains $\lceil n/2 \rceil$ (or $\lfloor n/2 \rfloor$) 1's.

Balanced codes are useful for the detection of unidirectional errors, that is, errors that can transform 0 into 1 but not 1 into 0 (or vice versa). Since each error changes the number of 1's in the codeword, balanced codes are *all-unidirectional error detecting*. Recent results show that many faults in VLSI circuits produce unidirectional errors. All-unidirectional error correcting codes are also useful for preventing subversions of the information recorded on 'write-once' memories, like digital optical disks.

Knuth first afforded the problem of constructing balanced binary codes having efficiently computable encoding and decoding functions and proposed both sequential and parallel efficient encoding/decoding schemes. Subsequently Al-Bassan and Bose improved Knuth's sequential encoding scheme. They show that it is possible to encode all binary sequences of length k with a balanced code of length $\kappa + p$ if $k \le 2^{p+1} - p - 2$.

In this paper we improve the construction proposed by Al-Bassan and Bose, that is, for a number p of parity bits and a number k of information bits, where

$$k = 2^{p+1} - 2x - 4 \quad \text{if } p \text{ is even, and} \quad k = 2^{p+1} - 2x - 5 \quad \text{if } p \text{ is odd,}$$

we prove that, if x satisfies

$$\tfrac{1}{2} \sum_{i=\lfloor p/2 \rfloor - x}^{\lceil p/2 \rceil + x} \binom{p}{i} \ge \sum_{i=1}^{\lfloor p/2 \rfloor - x - 1} i \binom{p}{\lfloor \frac{p}{2} \rfloor - x - 1 - i},$$

it is possible to encode all k-bits information sequences into a balanced code of length $k + p$. Relation (1) is satisfied at least for $x \ge \sqrt{p}$. Evaluations show that there are values of x less than $\sqrt{p}$ that satisfy relation (1) and suggest that a better bound might exist.

The encoding and decoding procedures are based on the sequential complementation of bits and have time and space complexity $O(k)$.

Decoding of Algebraic Geometry Codes

J. Justesen, K. J. Larsen, H. Elbrønd Jensen, and T. Høholdt *Institute of Circuit Theory and Telecommunication and Mathematical Institute, the Technical University of Denmark, DK-2800 Lyngby, Denmark*

Based on our earlier results on the decoding of codes from algebraic geometry we present some improvements of the algorithm. For a code of length n and designed distance d, complexity is $0(n^2)$ and most error patterns of weight $< d/2$ are corrected. We present a class of error patterns which was not corrected by previous algorithms, and discuss decoding methods for these cases.

Investigation on a New Class of Bilateral-Checking Codes

Jin Fan, Fan Pingzhi, and Chen Zhi *Dept. of Computer Science & Engineering, Southwestern Jiaotong University, Emei, Sichuan 614202, People's Republic of China*

Based on the Block Design theorem, a new kind of Bilateral-Checking Code is suggested in this paper. Just as its name sounds, "Bilateral-Checking" means that all the information bits are checked by the check bits, and conversely, all the check bits are checked by the information bits.

We show that under some given restrictions, Bilateral-Checking Codes can easily be derived by means of a symmetric balanced incomplete block design. Since this new class of linear codes is majority-logic decodable, a rather high decoding speed may thus be expected.

SESSION ThP7

ERROR-CORRECTING CODES II

A Family of Efficient Burst-Correcting Array Codes

Mario Blaum *IBM Research Division, Almaden Research Center, San Jose, CA 95120*

We present a family of binary burst correcting array codes defined as follows: consider an $n_1 \times n_2$ array with $n_1 = 4u + v + 2$ and $n_2 = 6u + 2v + 5$, $u \geq 1$, $v \geq 0$, $v \neq 1$, each row and column having even parity. The bits are read diagonally starting from the upper-left corner. The columns are viewed cyclically, i.e., the array is a cylinder. If one diagonal has been read out, proceed with the second diagonal preceding it.

We prove that codes of this type can correct any burst of length up to n_1. The burst-correcting efficiency of this family tends to 4/5 as $u \to \infty$. As a comparison, the burst-correcting efficiency of other families of array codes tends to 2/3. the same is true for Fire codes.

We also present a simple decoding algorithm for the codes defined above.

A New Burst and Random Error Correcting Code: The Projection Code

Gary R. Lomp and Donald L. Schilling *SCS Telecom, Inc., Port Washington, NY 11050, and City College, Department of Electrical Engineering, City College of New York, New York, NY 10031*

A new coding technique, capable of detecting and correcting both random errors and burst errors is described. The technique is based on an extension of majority logic coding, and in fact provides an effective design methodology for constructing such codes. Decoding is accomplished by a simple iterative scheme, wherein the number of errors is reduced significantly with each iteration. Both the encoder and decoder are easily implemented as either a block code or convolutional code. Binary and nonbinary codes are available and erasure decoding is readily performed. Bit rates in excess of 100 Mbs are readily achieved using programmable gate arrays and 2400b/s encoding and decoding can be achieved using a microprocessor based codec.

The decoding algorithm is capable of correcting many error patterns of weight in excess of the guaranteed minimum correction capability. It is shown how this feature may be effectively exploited in a hybrid code design that employs cyclic coding in a natural fashion within the present code. The resulting hybrid code has increased random error correction capability while retaining the burst error correction capability.

This paper contains a discussion of the encoding and decoding processes and performance data as estimated analytically and determined by simulation. Some details of the code structure and design of specific codes are also given.

Burst Asymmetric Error Correcting Codes

Seungjin Park and Bella Bose *Department of Computer Science, Oregon State University, Corvallis, OR 97331-3902*

Codes capable of correcting burst asymmetric errors are described. The proposed code needs approximately $b + \log_2 k + 1/2 \log_2 \log_2 k$ check bits to correct a burst of b asymmetric errors, where k is the number of information bits. Thus when $\log_2 k + 1/2 \log_2 \log_2 k < b$, the proposed code is better than any burst symmetric error correcting code. The optimality of the codes is also considered.

Burst-Error-Correcting and Detecting Codes

Kumar N. Sivarajan, Robert J. McEliece, and Henk C. A. van Tilborg *California Institute of Technology, Mail Code 116-81, Pasadena, CA 91125*

In this paper, we present several recent results in the area of simultaneous burst-correcting and detecting codes. We have obtained bounds similar to those of Reiger and Abramson for these codes, determined the true burst-correction and detection capabilities of Gilbert codes and developed efficient algorithms to determine the burst-correcting and detecting limits of cyclic codes. We also observe that many of the "good" burst-error correcting codes have very little burst-detecting ability. On the other hand, Fire codes which have strong burst-correction and detection capabilities may prove to be optimal.

An Application of Error-Correction Coding to Semiconductor Memories

C.P.M.J. Baggen *Philips Research Laboratories, P.O. Box 80.000, 5600JA Eindhoven, The Netherlands*

In this paper, we introduce a particular concept of ECC that is tailored to semiconductor memories. The main features of the considered class of ECC schemes are the regular repetitive structures which facilitate design, and the short decoding delays. Furthermore, the choice of the decoder site allows the definition of long codewords, which are necessary for high-rate codes.

The basic idea of the codes can be understood by considering them as convolutional codes, where the time coordinate has been replaced by the space (column) coordinate. We discuss a particular class of codes which appear to be equivalent to codes proposed by Wyner and Ash.

The considered class of codes appear to offer the required combination of low complexity, small decoding delay and high rate.

On the Correcting Capabilities of Product Codes

L.M.G.M. Tolhuizen and C.P.M.J. Baggen *Philips Research Laboratories, P.O. Box 80.000, 5600JA Eindhoven, The Netherlands*

In this paper we show that a product code is much more powerful than is generally expected. Despite the poor minimum distance, product codes may still offer a good performance, even with relatively simple decoding algorithms because of the low number of low weight codewords. Firstly, we classify and count all codewords of weight less than $d_r d_c + \max(d_r, d_c)$, where d_r and d_c refer to the distances of the row code and column code respectively. This leads to an upper bound on the number of low weight error patterns that a nearest neighbor decoder does not necessarily decode correctly. Secondly, we present a class of error patterns which have the all-zero word as closest codeword. This class suggests decoding possibilities beyond those already known for the simultaneous correction of burst errors and random errors.

ECC for Multi-Valued Random Access Memories

Rodney M. Goodman and Masahiro Sayano *Department of Electrical Engineering, 116-81, California Institute of Technology, Pasadena, CA 91125*

A generalization of simple codes for decoding approximate errors, that is, those errors which result in values which are approximately that of the true values, will be introduced. This situation arises when digital information is stored in analog levels and is corrupted only slightly in encoding or decoding, as in the case of multi-valued RAM. Such codes become important if data is to remain error-free as more bits are encoded in analog RAM cells and the size of the discrete steps which distinguish one digital data block from another decreases. In particular, errors of this type tend to occur in phased bursts due to timing or reference value errors, so phased burst errors must be decodable, while to maintain high data density and access rates the codes must be high-rate and easily encoded and decoded.

Upper and Lower Bounds on the Error Performance of Punctured Convolutional Codes

Guy Bégin and David Haccoun *Department of Electrical Engineering, Ecole Polytechnique de Montréal, P.O. Box 6079, station "A", Montréal (Quebec) Canada H3C 3A7*

High-rate punctured convolutional codes are derived from low-rate convolutional codes by the periodic elimination of code symbols after encoding according to a perforation pattern. Powerful punctured codes may be selected on the basis of bounds on the achievable error performance computed from the resulting characteristics (free distance, weight spectrum, distance profile, etc.) of the codes. Since the characteristics of a punctured code cannot be directly related to those of its original low-rate code, searching for powerful high-rate punctured convolutional codes involves a new evaluation of the characteristics for each combination of original code and perforation pattern. Likewise, ascertaining the error performance of a punctured code for a given coding rate R and memory M generally involves the full determination of its characteristics.

Fortunately, upper bounds relating the original and punctured weights of remerging paths may be derived. Using these relations, an upper bound on the error performance of a punctured code may be computed from the characteristics of the original code and the coding rate. On the other hand, using a completely different approach, a lower bound on the performances of punctured codes may also be derived. Based on the similarity that exists between perforation and erasures on an erasure channel, the performances of punctured codes with a random perforation pattern may be evaluated. Using classical arguments, these performances lower bound those of the best punctured codes. Comparisons between the lower and upper bounds and Viterbi's upper bound on error performance confirm the validity of the proposed approach for ascertaining the achievable error performance of punctured codes.

PLENARY SESSION

Friday, 8 - 8:50 a.m.

R. L. Dobrushin, *Institute for Problems of Information Transmission, Academy of Sciences of the U.S.S.R., 19 Ermolovoy St., Moscow, GSP-4, 101447, USSR*

TECHNICAL SESSIONS

Friday, 9 a.m. - 12 m.

SESSION FA1a

MAGNETIC RECORDING

Performance of Equalizers in Digital Magnetic Recording Channels

John G. Proakis and Dennis Tyner *Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA 02115*

The performance of linear, decision-feedback, and MLSE (maximum-likelihood sequence estimation) equalization techniques is evaluated for a digital magnetic recording channel. Both uncoded and run-length-limited (RLL) coded modulation are considered in the evaluation of the equalizer performance. Also evaluated is the performance of linear and decision-feedback equalizers which shape the desired signal pulse at the output of the equalizer into one of a family of partial response systems.

The Capacity Region for Write Unidirectional Memory Codes over Arbitrary Alphabets

W. M. C. J. van Overveld *Dept. of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Binary Write Unidirectional Memory (WUM) codes were designed for the rewritable optical disk. Writing is done in alternating 0- and 1-cycles: in a c-cycle, only c's can be written ($c \in \{0, 1\}$). Only the encoder knows the previous state of the disk.

We generalize this to WUM codes over an arbitrary alphabet $\{0, 1, ..., q-1\}$, where a c-cycle is followed by a $(c+1 \bmod q)$-cycle. With $M_c :=$ the number of messages that can be written in a c-cycle and $N :=$ the block length, we define $R_c := \log(M_c)/N$ and $\mathbf{C}_q :=$ the capacity region, i.e., the set of all achievable rate tuples $(R_0, R_1, ... R_{q-1})$. We derive the expression for $\mathbf{C}_q$.

For the symmetrical case, we find an elegant formula for the maximum C such that $(C, C, ..., C) \in \mathbf{C}_q$: we can write C in terms of the largest real zero of a polynomial. We also prove that symmetrical rate $R = (q-1)/q$ can be achieved with an easy code construction.

On the Capacity of the Bit-Shift Magnetic Recording Channel

S. Shamai (Shitz) and E. Zehavi *Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*

The bit shift (peak shift) effect is often encountered in magnetic recording and it rises mainly due to the effect of noise and read-out impairments and clock jittering. We investigate the degrading effect of this phenomenon on the capacity of a binary (d,k) coded system, commonly used in magnetic recording. The bit shift channel is best formulated in terms of phrase lengths. A phrase starts with none, one or more '0' and terminates with a single '1', thus: any binary {'0', '1'} sequence is uniquely decomposable into a concatenated sequence of phrases. The shift effect, restricted here to no more than a single channel bit position, either shrinks the input phrase length or expands it by 1. We restrict our discussion to $d \geq 2$, thus additional phrases are neither generated nor existing phrases phrases are destroyed. Keeping track of the inherent correlation of consecutive, bit shift affected, phrase lengths we observe that:

$$y_i = x_i + \varepsilon_i - \varepsilon_{i-1},$$

where x_i stands for the i-th channel input phrase length and any y_i is the corresponding channel output. The i.i.d. ternary random variable ε_i taking on (-1,0,+1) values designates whether a bit shift has occurred ($\varepsilon_i = \pm 1$) or not ($\varepsilon_i = 0$).

An increasingly tight and easily calculable sequence of upper and lower bounds on capacity (normalized per channel bit input)

$$C = \lim_{N\to\infty} \sup E(\underline{X}^N)^{-1} I(\underline{Y}^N : \underline{X}^N)$$

is derived. Here $\underline{X}^N, \underline{Y}^N$ stand respectively for input and output vectors of N consecutive phrase components, $I(\cdot,\cdot)$ and $H(\cdot)$ stand respectively for the mutual information and entropy functionals, E is the statistical average operator and the supremum is taken over $P(\underline{X}^N) \in P(d,k)$ - the probability measure of $\underline{X}^N$ satisfying the (d,k) constraint, that is: $x_i \in (d+1, d+2,\ldots,k+1)$ $i = 1,2,\ldots,N$. The capacity of a concatenated scheme in which the bit shift channel is followed by a binary symmetric channel is also addressed and certain upper and lower bounds are reported.

Spectral Null Codes

K. A. Schouhamer Immink *Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands*

Often, the servo position information of magnetic tape or disk recorders is recorded as low-frequency components, usually called pilot tracking tones. To circumvent interaction during read-out between the pilot tones and the user information, user information is often encoded in such a way that the power spectral density function of the encoded stream vanishes at the pilot tone frequencies. Binary codes giving rise to a spectral null at an arbitrary frequency are used to provide space for the allocation of auxiliary pilot tones.

This paper deals with encoding methods in which binary data are mapped into constrained binary sequences for shaping the spectrum. We compute the rate and power spectral density function of memoryless codes that exhibit spectral nulls. The relationship between code redundancy and spectral notch width is quantified with a parameter called the sum variance.

SESSION FA1b

INFORMATION THEORY APPLICATIONS

Necklace Properties of Shuffle-Exchange Graphs

S. Song and E. Shwedyk *Department of Electrical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2N2*

Shuffle-exchange (SE) graphs, originally used as an interconnection methodology for parallel computations, have been proposed as a VLSI layout method for the implementation of the Viterbi algorithm and the Fast Fourier transform. Some asymptotically optimum VLSI-layouts are based on SE graphs.

A key part of the layout technique is partitioning the graphs's shuffle edges into necklaces. This paper presents results which show:

a) under what conditions a particular shuffle-exchange network contains only full necklaces (excluding self-loops).

b) when the graph has degenerate (nonfull) necklaces and an enumeration of their lengths.

c) a systematic method of listing all necklaces.

It is shown that counting necklaces is independent of memory length and so enables the table construction which lists the number of binary necklaces for SE graphs of various memory length.

Selection-Based Locally Connected VLSI Architectures for the (M,L) Algorithm

E.M. Leiby III and Seshadri Mohan *ECE Dept., Clarkson University, Potsdam, NY 13676*

The (M, L) algorithm has found applications in speech and image coding, decoding of convolutional codes, cpm codes, and trellis codes, and in the decoding of sequences received over intersymbol interference channels. With its ever increasing applications, and to cope with the requirement of high data rates, there is an urgent need to design architectures for parallel processing suitable for VLSI implementation. Here we propose an SIMD mesh- and a tree-based architectures for the algorithm and show how the algorithm may be implemented on these. We compare these to two other architectures proposed in the literature, one due to Mohan and Sood and the other due to Simmons.

Using Decision Trees for Noiseless Compression

P. A. Chou *AT&T Bell Laboratories, Rm 2C-479, 600 Mountain Avenue, Murray Hill, NJ 07974*

We use a binary decision tree for noiselessly compressing a finite-alphabet stationary stochastic process $\{X_n\}$, by testing, at each internal node, one of the k previous symbols $X_{n-1},\ldots,X_{n-k}$, and by outputting, at each leaf, the distribution of the current symbol X_n. Since the output distribution of X_n depends on the previous k symbols, the tree models the conditional distribution of X_n, and can be used to drive a conditional entropy coder for noiseless compression. We build the decision tree by recursively applying, at each node, a recently developed algorithm for partitioning the alphabet of a discrete random variable X into two bins, such that the conditional entropy of a related random variable Y (given the bin) is minimized. Thus the tree is explicitly designed to minimize the conditional entropy of X_n, given the leaves of the tree. The partitioning algorithm is equivalent to the K-means algorithm for pattern clustering, or to the generalized Lloyd algorithm for vector quantizer design, but uses Kullback's information divergence, or relative entropy, as the distortion measure. Using this method, we reduce the bit rate of an image coder by 46 percent.

Lower Bounds on the Capacity of Asymptotically Good Spherical Codes in the Gaussian Channel

Dejan E. Lazić *Universität Karlsruhe, Institut für Algorithmen und Kognitive Systeme, D-7500, Karlsruhe 1, Fed. Rep. of Germany*

The capacity of a sequence of a specific family of codes with the fixed bit code rate is defined as the least upper bound on permissible bit code rates, i.e. on bit code rates for which probability of error decreases to zero as the number of code dimensions grows beyond all bounds. For the family of asymptotically good spherical (i.e., equal energy) codes, a new lower bound on the code capacity is derived, and compared to those found in the literature. The new bound is greater than zero for signal-to-noise ratios below 0 dB, which was not a feature of the previous ones.

SESSION FA2

PATTERN RECOGNITION

A Reformulation of the EM Algorithm for Hidden Markov Model Parameter Estimation

M. Ostendorf and J. R. Rohlicek *ECS Department, Boston University, Boston, MA 02215, and BBN Systems and Technology Inc., Cambridge, MA 02138*

Hidden Markov models (HMM) are used in pattern recognition to represent time-varying signals, such as speech. A hidden Markov model consists of a discrete-time Markov process representing an unobserved state sequence, with random observations which are conditionally independent given the state sequence. The focus of this paper is on the parameterization of the joint likelihood of the state and observation sequences. In particular, we present a parameterization which is not typically used and show that the maximum likelihood parameter estimate is related to a maximum mutual information criterion between observations and states. This suggests a method for jointly estimating the parameters of a discrete hidden Markov model and an observation vector quantizer. A tree quantizer design algorithm based on maximizing the mutual information between the observations and the model states is presented. This quantizer has an additional advantage in that it is well-suited to the use of non-homogeneous features.

A Two-Dimensional Maximum Likelihood Approach for Image Edge Detection

Jack Koplowitz and Xiaobing Lee *Department of Electrical and Computer Engineering, Clarkson University, Potsdam, NY 13676, and SBC Technology Resources Inc., 550 Maryville Center Dr., St. Louis, MO 63141*

Traditionally, the problem of edge detection in images has been to locate those pixels which contain an edge, i.e., a jump in the intensity level. Recently a number of researchers have turned their attention to obtaining subpixel accuracy in estimating edge location. This work has generally focused on matching a parametric surface to the region in the image where an edge occurs. Here a two-dimensional maximum likelihood approach is used to obtain the most likely position and orientation of the edge. Comparisons show that the maximum likelihood approach gives considerably less error in locating the position of the edge over a wide range of noise conditions, as well as blurring which may be unknown to the estimator.

Approximate String Embedding in a Labeled Graph

San Wei Sun and S. Y. Kung *Telecommunication Laboratories, Ministry of Communications, P.O. Box 71, Chung-Li, Taiwan, R.O.C., and Department of Electrical Engineering, Princeton University, Princeton, NJ 08544*

This paper presents a *dynamic programming* to embed a given string within a directed, connected and labeled graph with flexible error correction metric. Common error correction operations, namely, omissions, insertions, substitutions, and the reversals, are tackled. The proposed algorithms will find an optimal path in such a graph with a minimal cost of corrections. Two versions of dynamic programming algorithms are devised to find the optimal paths. The embedding problems with omission, insertion, and substitution operations are modeled by a trellis diagram, then a Viterbi algorithm is applied to find the best path in the trellis state transition diagram. For problems with omission, insertion, substitution and reversal operations, a linear embedding procedure is devised to rearrange a graph in a linear form, then a general dynamic programming algorithm is proposed to find the optimal path. The scheme is applied to an *On-Line Chinese Character Recognition* problem with a variety of writing scripts for a character, the experimental results have verified the feasibility and effectivity of this approach.

Description lengths of data and model parameters are calculated on the basis of these distributions. Next, we apply the MDL criterion to selection of the best decision list. That is, for any given data, the decision list selected as the best will be that which requires the least total description length for the encoding of its model parameters and of the data observed through it. Here model parameters are estimated on the basis of their hierarchical structure. Finally, the optimality of the inferred decision lists is proved in terms of their convergence and the convergence speed with which for any $\varepsilon > 0$, $Prob\,[D(P^* \,\|\, Pht_N) \geq \varepsilon]$ goes to zero as N becomes large. Here $D(P^* \,\|\, Pht_N)$ is the Kullback-Leibler distance between a true data distribution P^* and a data distribution Pht_N defined by the decision list inferred from i.i.d. N data. *Prob* is N direct product of the probability measure P^*. The optimality of the inferred decision lists is also experimentally discussed.

Hidden-Markov-Model Based Optical-Character Recognition - a Novel Approach

Bor-Shenn Jeng, Fu-Hua Liu, San-Wei Sun, and Tiei-Min Wu *Telecommunication Laboratories, Ministry of Communications, P.O. Box 71, Chung-Li, Taiwan, R.O.C.*

In this paper, we present the application of the Hidden Markov Model as a learning and recognition model for multi-font Chinese character recognition. This doubly stochastic process encodes the distortion and similarity among patterns of a class thru a stochastic and evaluating approach. A simple feature extraction is employed in our experiments, which is the histogram of the projected profiles. A character class is modeled by a 7-state Hidden Markov model, and the recognition rate for 1000 multi-font (five styles) character classes reaches 98% in average for inside test and 91% in average for outside test. This is a novel approach and is rarely applied to resolve character recognition problems in the open literature.

On-Line Recognition of Handwritten Chinese Characters using Nonlinear String Matching Method

Chang-Keng Lin and Bor-Shenn Jeng *Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, P.O. Box 71, Chung-Li, Taiwan, R.O.C.*

This paper proposes an on-line handwritten Chinese character recognition system using a nonlinear matching method. The character to be recognized can be stroke-order and stroke-number free, tolerant for combined strokes, size flexible, but within the constraint of normal hand-writing. In the first recognizer of the algorithm the finite-state matching mechanism is used to generate a stroke string for the input test character. In the second recognizer, a windowed-dynamic-programming (WDP) based nonlinear matching method is used to perform recognition with the stroke-string features. The alternative linear matching method and its experimental results are also submitted for comparison and analysis. Reference patterns have been generated for 2000 Chinese characters with stroke numbers ranging from 1 to 27. The recognition results are based upon the 1800 handwritten characters by 10 people. The obtained recognition rate is 94.5%, and the cumulative classification rate of choosing fourth most similar characters is up to 99%. The multiple dp matching method and subarea matching method implemented in the simulation system to achieve rather high speed performance are also presented.

On the Image Recovering from Moment Descriptors

Miroslaw Pawlak *Dept. of Electrical Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2*

The problem of reconstruction of an image from noisy data by the method of moments is examined. The set of orthogonal moments based on Legendre polynomials is employed. A general class of signal-dependent noise models are taken into account.

Classification with a Reduced Complexity Nearest Neighbor Algorithm

Tamás Linder and Gábor Lugosi *Institute for Communication Electronics, Technical University of Budapest, Budapest Hungary*

In 1986 Vidal presented an algorithm, which on the basis of some simple geometric considerations, finds the nearest neighbor in a metric space much faster than the exhaustive search. He illustrated the effectiveness of the algorithm by simulation results. Our paper analyses the speed of the algorithm when a certain probability model is assumed. We prove, that if the metric space is the $\mathbb{R}^n$ with the Euclidean metric, then under some suitable conditions, the number of distance calculations required for finding the nearest neighbor tends to be negligible compared to that of the exhaustive search as the number of sample points tends to infinity. Closely related to the nearest neighbor algorithm we derive a classification method with asymptotically the same error probability as that of the nearest neighbor decision, for which the number of necessary distance calculations depends only on the dimension of the space rather than on the number of the sample points.

A Parameter Estimation Algorithm for Speech Recognition to Maximize "State Optimized Joint Likelihood"

G. Lugosi and A. Faragó *Institute of Communication Electronics, Technical University of Budapest, H-1111 Stoczek u.2, Budapest, Hungary*

A central problems concerning speech recognition using statistical methods is the estimation of the probability distributions of the Hidden Markov Models which are supposed to describe the time-varying properties of the speech signal. The distributions are determined by finite number of parameters and the estimation is based on the maximization of an objective function of these parameters.

Most commonly, the object of the maximization is to obtain maximum likelihood or conditional maximum likelihood estimation. Existing algorithms, such as Baum-Welch reestimation or its analog to conditional maximum likelihood estimation provide convergence to a local optimum of these objective functions.

State optimized joint likelihood is another objective function for which such algorithms exist (Jelinek's "Viterbi extraction" or Rabiner, Wilpon and Juang's "Segmental k-means" algorithm.

In this paper we present an algorithm which finds the globally optimal parameters--in the sense of maximizing state optimized joint likelihood--for the class of left-to-right Hidden Markov Models, if an observation sequence is given.

Inferring Optimal Decision Lists from Stochastic Data Using the Minimum Description Length Criterion

Kenji Yamanishi *C&C Information Technology Research Laboratories, NEC Corporation, 1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki, 213, Japan*

Inferring classification rules from stochastic data is an important issue in machine learning. This paper, in terms of information theory, presents a method for inferring classification rules using Rissanen's Minimum Description Length (MDL) criterion and evaluates the performance of the inferred rules.

We deal with most specifically the classification rules called "decision lists", which were first proposed by Rivest. These specific formulae take the form of linearly sequences of "if - then - else if ..." type decisions. Although Rivest formulated the decision lists as deterministic rules, we can consider them linear sequences of "if - then - (with probability p) else if ..." type stochastic decisions, when dealing with data from stochastic information sources. Such classification rules define probability distributions of data, and each rule itself can be regarded as a probabilistic model of an information source. Inferring decision lists is equivalent to estimation of these probabilistic models.

In this paper, first, we formulate the probability distributions of data with respect to individual decision lists. Here each decision list has model parameters forming a hierarchical structure. We further formulate the prior distributions of those model parameters according to this hierarchical structure.

An asymptotic expansion for the global reconstruction error is established. This reveals mutual relationships between a number of moments, the image smoothness, sampling rate and noise model characteristics.

The problem of an automatic (data-driven) selection of an optimal number of moments is studied. This is accomplished with the help of cross-validation techniques.

SESSION FA3

SIGNAL PROCESSING III

Performance Analysis of the Constrained LMS Algorithms with Uncorrelated Gaussian Data

Abraham Krieger *Orincon Corporation, 9363 Towne Centre Drive, San Diego, CA 92121*

The convergence properties of a constrained adaptive filtering algorithm are established for uncorrelated Gaussian input data. The constraints are in the form of bounded sets in which the filter's coefficients must lie. Two constraint sets in R^n are considered: A bounded hypercube with all edges equal to B_1, and a bounded hypersphere with radius equal to B_2. A bound on the mean-square-deviation (MSD) of the estimate of the filter's coefficients from its optimal one was obtained in earlier work for dependent input processes. This bound is a second-order polynomial of the parameters B_i, $i = 1, 2$. The goal of this work is to show that when the input data is an uncorrelated Gaussian process the bound is a bounded function of B_i rather than a second-order polynomial.

A Geometrical Approach to Multiple-Channel Detection

Douglas Cochran and Herbert Gish *Department of Electrical and Computer Engineering, Arizona State University, Tempe, AZ 85287, and BBN Systems and Technologies Corporation, 10 Moulton Street, Cambridge, MA 02138*

The *Generalized Coherence (GC)* estimate, a recently introduced detection statistic that forms the basis of a non-parametric test for common but unknown signal on several noisy channels, is shown to be geometrical in nature. This perspective is used to show that the GC estimate is the unique function of $M \geq 2$ complex sample sequences having a desirable set of properties. Its distribution function is derived under the H_0 assumption that the sample sequences representing filtered data from M noisy channels are drawn from independent Gaussian processes. This is accomplished by exploiting the interpretation of the MSC estimate as the volume of a polytope in (real) $2M$-dimensional space. The properties and performance of the GC estimate as a detection statistic are also discussed. Detection thresholds are derived corresponding to different false alarm probabilities for various numbers of channels and sequence lengths. The performance of the GC detector as a function of the signal-to-noise ratios on the noisy channels is evaluated by simulation and compared to performance data for other multiple-channel detection schemes.

A Fast and Effective Algorithm for Sinusoidal Frequency Estimation

S. F. Hsieh, K. J. R. Liu, and K. Yao *Electrical Engineering Department, University of California, Los Angeles, Los Angeles, CA 90024-1594*

Recently, a myriad of researches have been focused on SVD and eigen-decomposition analysis of multiple sinusoidal frequency estimations, which require iterative massive computations, hence are quite intractable for real time applications. Here we propose a direct method, involving only a fixed number of operations, called *truncated QR*(TQR), which is a simplified modification of a SVD-based FBLP method developed by Tufts and Kumaresan. Since the QR method is amenable to VLSI (e.g., systolic array) implementation, a real-time processing for identifying multi-harmonics is thus feasible under stationary and time-varying conditions.

A FBLP model is used in the solution of the rank-reduced LS problem. Then an AR pseudo-spectrum is plotted, where the peaks yield the locations of the harmonics embedded in the noise perturbed data samples observed over a finite duration. Our method shares the benefits of identifying narrowly spaced harmonics as well as suppressing spurious frequency peaks such as in the truncated SVD(TSVD) approach. However, the computational cost has been greatly reduced. Computer simulations are provided for comparisons to other known approaches and show that the TQR method is a promising candidate for the efficient retrieval of clustered harmonics in real time applications.

Time-Frequency Filtering and Synthesis from Convex Projections

Langford B. White *Electronics Research Laboratory, Defence Science & Technology Organization, P.O. Box 1600, Salisbury, S.A., 5108 Australia*

This paper describes the application of the theory of projections onto convex sets (POCS) to time-frequency filtering and synthesis problems. We show that the class of Wigner-Ville Distributions (WVD) of L_2 signals form the boundary of a closed convex subset of $L_2(\mathbb{R}^2)$. This result is obtained by considering the convex set of states on the Heisenberg group, of which the ambiguity functions (AF) form the extreme points. The form of the projection onto the set of WVD's is deduced. Various linear and nonlinear filtering operations are incorporated by formulation as convex projections. An example algorithm for simultaneous time-frequency filtering and synthesis is suggested.

Modifying Real Convolutional Codes for Protecting Digital Filtering Systems

Robert Redinbo and Bernhard Zagar *Dept. of Electrical Engineering and Computer Science, University of California, Davis, Davis, CA 95616, and Institut für Allgemeiner Elektrotechnik und Elektrische Messtechnik, Technical University of Graz, A-8010, Graz, Austria*

Digital filters when implemented with very dense high-speed electronic devices are susceptible to both temporary and permanent failures, not easily protected by conventional fault-tolerant computer design principles. A new method is presented for protecting the overall realization against both hard and soft errors at the data sample level using the error-detecting properties of real convolutional codes. The normal filter system is surrounded with parallel parity channels defined by a real systematic convolutional code. Erroneous behavior is detected by comparing externally the calculated and regenerated parity samples. A rate $\frac{k}{n}$ real convolutional code produces $n-k$) parity samples for every k input samples causing the parity channels to operate at a rate decimated by k. Significant complexity reductions are possible by modifying the code structure, without loss of error protection, yielding simplified parity channels with finite impulse response (FIR) structures operating at rate decimated by k. The theory and practice of the code modification procedure depend on an annihilator subspace associated with the columns of a code segment matrix. This subspace is affiliated with the dual code description. Among other things, the parity channels are redesigned to cancel poles associated with the original transfer function. A set of homogeneous equations determine the class of solutions by interrelating dependent and independent choices for the modified parity channels' FIR parameters.

First Order Error PDF for Nonlinear Stochastic Filters

Carlos A. C. Belo and José M. F. Moura *Dept. of Elect. and Comp. Eng., Carnegie Mellon University, Pittsburgh, PA 15213*

The paper studies the characterization of the error for nonlinear estimation problems. In these problems, filters are designed to provide estimates of an underlying unobserved process (the message of interest) from noisy measurements of a nonlinear function of the message. An outstanding unresolved question is the performance evaluation of the optimal as well as suboptimal nonlinear filters. In this work, we consider directly the first order characterization of the error process $p(\varepsilon)$ of any nonlinear filter F. It turns out that this function is obtained as an expectation of a translated version of the a posteriori pdf.

Let x be the signal to be estimated and Z the available observations. If $p_{ap}(Z) = p_{x|Z}(X|Z)$ is the conditional density of x given Z then the filter F^1 that computes the estimate $\hat{x}^1 = \mathsf{F}^1(Z)$, exhibits an error $\varepsilon = x - \hat{x}^1$ with probability density function $p_{\varepsilon}^{\mathsf{F}^1}(X) = E[p_{ap}(X + \hat{x}^1)] = E[p_{\varepsilon|Z}^{\mathsf{F}^1}(X)]$, where $E[\cdot]$ is the expectation operator taken over the underlying joint probability space.

The technique has been applied to the absolute phase demodulation problem. The message is a two-dimensional signal that phase modulates a carrier. The error probability function has been computed for several nonlinear phase estimators using the method described above. Movies of the error pdf's for these filters will be presented.

Fast "Modified Triangular Factorization" of Hermitian Toeplitz and Quasi-Toeplitz Matrices with Arbitrary Rank Profile

Debajyoti Pal and Thomas Kailath *Information Systems Laboratory, Stanford University, Stanford, CA 94305*

We present a fast procedure for computing a "modified triangular factorization" of certain structured matrices in $O(n^2)$ operations. This family of structured Hermitian matrices includes Toeplitz and Quasi-Toeplitz (matrices with certain hidden Toeplitz structure).

The derivation of this procedure is based upon a function-theoretic interpretation of the so-called "displacement structure" exhibited by these matrices. This interpretation has been used by Lev-Ari and Kailath to derive fast triangular factorization procedures for the "strongly regular" Hermitian matrices with a so-called "generalized displacement structure." In this paper we have modified their procedure for "strongly regular" Toeplitz and Quasi-Toeplitz matrices to handle matrices with the same structures but having arbitrary rank profile. In case of "strong regularity" our new procedure produces the same result as Lev-Ari and Kailath's. If "strong regularity" is lacking then our procedure recursively produces a LDL^* factorization where D is a block diagonal matrix. Rank profile and inertia are also computed with no extra effort.

An Order Selection Rule for Rank Reduction in the Linear Statistical Model

Richard T. Behrens and Louis L. Scharf *Department of Electrical & Computer Engineering, University of Colorado, Boulder, CO, 80309*

A new order selection rule is proposed for rank reduction in the linear statistical model. Rank reduction decreases noise-induced estimator variance at the expense of increased estimator bias, but results in lower mean squared error. The new order selection rule is to find the maximum likelihood estimate of the mean squared error between the true signal and a reduced rank signal estimate, and choose the rank for which this estimated mean squared error is smallest.

SESSION FA4

CODING THEORY VIII

Extremal Codes are Homogeneous

Vera Pless *Mathematics Department, University of Illinois at Chicago, Chicago, Illinois, 60680*

We show that extremal codes are homogeneous. This implies that each punctured code of an extremal code has the same weight distribution, which can be calculated directly from the weight distribution of the parent extremal code.

Extremal Codes of Length 40 and Automorphism of Order 5

V. Y. Yorgov and N. P. Ziapkov *Mathematics Department, Higher Pedagogical Institute, Shumen 9712, Bulgaria*

Doubly-even self-dual (DESD) codes exist only in lengths n that are multiple of 8. All such codes are completely determined up to length 32. Those of DESD codes that have the greatest possible weight $4\lfloor n/24 \rfloor + 4$ are called *extremal*. Many extremal DESD codes of length 40 have been constructed recently via Jadamard matrices, some of them having a very small or trivial automorphism group. It is known that the only primes that can divide the group order of an extremal DESD code of length 40 are 19, 7, 5, 3, and 2. All such codes having automorphism of order 19 or 7 are known. Using the well-developed techniques for constructing self-dual codes via automorphisms, we obtain a complete list of all extremal DESD codes of length 40 having automorphism of order 5.

Equivalences of Binary Irreducible Goppa Codes

Hermann J. Helgert *Department of Electrical Engineering and Computer Science, The George Washington University, Washington, DC 20052*

We consider the class of binary irreducible Goppa codes of length 2^τ generated by the elements of $GF(2^\tau)$ and the Goppa polynomial $g(z) = g_1(z)^{2^t}$, where $g_1(z)$ is a polynomial of degree μ with coefficients from $GF(2^\tau)$ that is irreducible over $GF(2^\tau)$ and t is a positive integer. It is known that among this class there exist codes that asymptotically satisfy the Varshamov-Gilbert bound on performance as τ becomes large, although the problem of analytically determining which $g(z)$ generates such codes has not been solved. It is therefore of interest to develop methods for reducing the scope and complexity of a computer-based evaluation of the weight spectra of these codes. Toward this end we show that certain linear transformations of the roots of $g_1(z)$ yield Goppa polynomials that generate codes with identical weight spectra. This property allows us to collect the set of all irreducible Goppa codes of given degree into equivalence classes, thereby considerably reducing the number of codes that must be evaluated. We also obtain a result that relates the size of an equivalence class to the number of conjugates of a certain invariant of the roots of $g_1(z)$ and derive upper and lower bounds on the number of equivalence classes for given τ, μ and t that are asymptotically tight as τ becomes large.

The Automorphism Group of the Kerdock Code

Claude Carlet *Universite' Paris VI, Laboratoire d'Informatique Theorique et Programmation (LITP), 4 place Jussieu, 75252 Paris Cedex 05, France*

We prove that the automorphism group of the Kerdock code of length 2^m, m even and $m \geq 6$ is the group of the following permutations on $GF(2^{m-1}) \times GF(2)$:

$$(x, \varepsilon) \in GF(2^{m-1}) \times GF(2) \rightarrow (ax^{2^i} + b, \varepsilon + \delta)$$

where a ranges over $GF(2^{m-1}) - \{0\}$, b ranges over $GF(2^{m-1})$, $\delta \in GF(2)$ and $i \in [0, m-2]$.

Thus, the Kerdock and Preparata codes of same length, which are known to be formally dual, have also the same automorphism group (W. Kantor has determined Preparata's one).

In the proof of this new result, we solve the principal problem by using the relation between the weight of an element of the Reed-Muller code of order 2 and the rank of its symplectic form. We use also and prove the following lemma:

Let Ψ be an automorphism of the linear space $G' = GF(2^{m-1})$, and f be a linear form on G' such that for all $(x,y) \in G'^2$ we have:

$$(\Psi(x)\Psi(y))^{2^{m-2}} + f(x)\Psi(y) + f(y)\Psi(x) + \Psi((xy)^{2^{m-2}}) = 0.$$

Then f is the null function and Ψ is of the type:

$$x \to ax^{2^i}, \quad i \in [0, m-2], \quad a \in G' - \{0\}.$$

There is No (24, 12, 10) Self-Dual Quaternary Code

Clement Lam and Vera Pless *Department of Computer Sciences, Concordia University, Montreal, Quebec, Canada H3G 1M8, and Department of Mathematics, Statistics and Computer Science, University of Ilinois at Chicago, Chicago, IL 60680*

The question of the existence of a (24, 12, 10) self-dual quaternary code has been a long-standing one. The largest minimum weight a self-dual quaternary $(n, n/2)$ code can have is $d = 2(n/6) + 2$ and a self-dual quaternary code attaining this minimal weight is called an extremal code. Extremal codes exist of all even lengths n less than or equal to 22. If an extremal quaternary code of length $n = 24$ exists, vectors of weights 10, 12 and 14 would hold 5-designs by the Assmus-Mattson theorem. We approached this problem by attempting to construct the generator matrix of this code directly. The conclusion we reached is that there is no such code.

The Carlitz-Uchiyama Bound and the Dual of the Melas Code

Gilles Lachaud *Equipe "Arithmetique et Theorie de L'Information", C.I.R.M., Luminy Case 916, F13288-Marseille Cedex 9, France*

The first part is devoted to the Carlitz-Uchiyama bound, which gives estimates on the weights of the words of the duals of BCH codes in characteristic two; as is well-known, its proof lies on some kind of Weil's inequality on exponential sums. We give here a bound for the duals of geometric BCH codes on curves of arbitrary genus on any prime field, i.e., for the subcodes over the prime field of geometric Goppa codes. To this end we give a sharpened bound for a quite general family of exponential sums along a curve. We thus give also some bounds on the number of solutions of the equation $Tr(f(x) = 0$, where f is a function define on a curve.

The second part is concerned with the study of a particular code in characteristic two, namely the dual of the Melas code that we call the Kloosterman code. The weights are connected to the classical Kloosterman sums. For this code we compute the weights exactly occurring, the multiplicity of weights in terms of class number of quadratic forms, and we show that in some sense these weights become equidistributed for a particular measure when q is large.

Exponential Sums and Goppa Codes

Carlos J. Moreno and Oscar Moreno *Baruch College and Graduate Center, CUNY, 17 Lexington Ave., Box 509, New York, NY 10010, and Department of Mathematics, University of Puerto Rico, Rio Piedras, Puerto Rico 00931* (40 min.)

Recently we obtained a sharpening of the Carlitz Uchiyam bound in characteristic two using Serre's improvements on Weil's estimate of the number of rational points on an algebraic curve over a finite field. In this talk we derive an improvement, also on characteristic two, on a theorem of Bombieri and

Weil. this has important applications in the theory of binary codes; for instance, we obtain immediate results concerning the minimum distance of the dual of an arbitrary subfield subcode of a Geometric Goppa. Furthermore, we estimate the covering radius, and we find the exact value for the minimum distance and the number of information symbols whenever the minimum distance is small in relation to the length of the code.

Research problem 12.1 in Mac Williams & Sloane "The Theory of Error Correcting Codes" asks for the true dimension and minimum distance of a classical Goppa Code. Using the above results for Geometric Goppa codes in the genus 0 case, we can solve the problem for the minimum distance of classical Goppa codes whenever the length of the code $n = 2^m$ satisfies a certain inequality on the degree of the Goppa polynomial. On the other hand, given that the approach in the first paragraph makes use of heavy number theory arguments, we will also derive an elementary proof of the improvement on the conditions of the theorem of E. Bombieri, which will give us a valid proof of the result in the case of classical Goppa codes. We use this improvement also to generalize a previous result on the minimum distance of the dual of a Goppa code.

SESSION FA5

ENTROPY

Entropy and Surface Entropy of Random Fields on Trees

Toby Berger and Zhongxing Ye *School of Electrical Engineering and Center for Applied Mathematics, Cornell University, Ithaca, NY 14853*

We study entropic aspects of several models of random fields on trees. For any shift invariant random field on a open Cayley tree, the entropy rate and surface entropy rate exist. In particular, for a Markov chain field the entropy rate has a simple form similar to that of a one-dimensional Markov chain. However, surface entropy rate and conditional surface entropy rate, which are trivial in the one-dimensional case, exhibit interesting behaviors when extended to Markov fields on trees. The above results are extended to any open Bethe tree on which each vertex has the same number of neighbors. We also consider random fields defined on closed trees with recursive structure. It is shown that the entropy rate exists for such random fields if they possess appropriate stationary properties.

The Entropy Power and Related Inequalities

Amir Dembo *Information Systems Laboratory, Stanford University, Stanford, CA 94305*

Two proofs of the Entropy Power Inequality (EPI) are presented. It follows from the convexity inequality $h(\sqrt{\lambda}\,X+\sqrt{1-\lambda}\,Y)\geq\lambda h(X)+(1-\lambda)h(Y)$, for independent random variables X, Y and $0\leq\lambda\leq 1$. Stam's proof of the EPI via implicit normal perturbations is simplified as explicit normal perturbations prove the convexity inequality.

A new proof of the convexity inequality via a limiting argument on Beckner's sharp version of Young's inequality is presented. A similar limiting argument on the converse of this inequality yields an analogue convexity inequality about Minkowski-sum of sets, which in turn results in the Brunn-Minkowski Inequality (BMI). Beckner's version of Young's inequality, and its converse, are restated in terms of Renyyi differential entropies, thus filling the gap between the BMI and EPI.

In addition to the isoperimetric inequalities that result from the EPI and BMI, a third isoperimetric inequality is derived from an inequality about Fisher informations. This isoperimetric inequality amounts to Costa's recent result on the concavity of the entropy power, thus supplying a short and elegant proof of the latter.

Graph Entropy and Convex Programming Dualities

Victor K. Wei *Bellcore, 435 South Street, Morristown, NJ 07960*

Shannon defined the (zero-error) capacity of a graph to be the maximum reliable transmission rate of a channel corresponding to the graph. Lovasz in 1979 derived the capacities of the pentagon and other graphs. Körner generalized the concept to include probabilistic considerations and defined the entropy of a graph. Recently, Csiszar, Körner, Lovasz, Marton, and Simonyi obtained a characterization of antiblocking pairs of polyhedra and of perfect graphs based on a certain strong splitting property of graph entropy.

Motivated by a resource allocation problem, Wei, and Monma, Schrijver, Todd, Wei studied a convex programming problem over acyclic directed graphs. Their results include various dualities of convex programs, approximate solution algorithms, and a generalization of the LYM property for partially ordered sets.

Here, we show that these two seemingly unrelated series of results can be unified. The entropy splitting property actually corresponds to a strong convex programming duality. With this approach, we obtain a simplified proof of the results in Csiszar, Körner, Lovasz, Marton, and Simonyi and generaliza-

tions. This may lead to a better understanding of the Shannon capacity of graphs and of several multi-user information theory problems.

Tight Upper Bounds on the Entropy Series

Renato M. Capocelli and Alfredo De Santis *Dipartimento di Matematica, Universitá di Roma, 00185 Roma, Italy, and IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York, 10598*

Two tight upper bounds on the entropy $H(P)$ of a countable random variable are provided. the first bound, expressed in terms of $E_p(\log)$, is

$$H(P) \le \mu E_p(\log) + \log \zeta(\mu),$$

where $\zeta(\mu)$ is the Riemann zeta function, μ is the unique solution of the equation $E_p(\log) = \Sigma_{k=1}^{\infty} \Delta(k)/k^{\mu}$; and $\Delta(k)$ is defined as $\ln p$ when k is a power of a prime p; 0 otherwise. The bound improves previous known upper bounds.

The second bound, expressed in terms of $E_p(\lfloor \log \rfloor)$, is

$$H(P) \le E_p(\lfloor \log \rfloor) + (E_p(\lfloor \log \rfloor + 1)\, \mathsf{H}\left[\frac{1}{E_p(\lfloor \log \rfloor) + 1}\right]$$

where $\mathsf{H}(p)$ is the Shannon function $-p \log p - (1-p)\log(1-p)$.

This latter bound allows to derive an easily computed approximation of the former and a bounded error estimate of it (the exact estimate would involve a complex evaluation of the Riemann zeta function)

$$\mu\, E_p(\log) + \log \zeta(\mu) = E_p(\log)\left[1 + \mathsf{H}\left(\frac{1}{E_p(\log)}\right)\right] + \alpha,$$

where α is a bounded function of $E_p(\log)$, i.e., $-1 \le \alpha \le \Delta_p - 1$, $\Delta_p \le 3$ if $1 \le E_p(\log) < \phi$ and $\Delta_p \le 1 + \mathsf{H}(1/E_p(\log))$, otherwise; and ϕ is the golden ratio $(1+\sqrt{5})/2$. (This work was partially supported by the Italian Ministry of Education, Project: Progetto ed Analisi di Algoritmi.)

When is Graph Entropy Additive? Or: Perfect Couples of Graphs

János Körner, Gábor Simonyi, and Zsolt Tuza *Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, P.O.B. 127, Hungary, and Computer and Automation Institute of the Hungarian Academy of Sciences, H-1111 Budapest, Kende u. 13-17, Hungary*

Graph entropy is an information theoretic functional on a graph and a probability distribution on its vertex set. fixing the probability distribution, it is sub-additive with respect to graph union. This property has been used to derive non-existence bounds in graph covering problems by several authors.

Here we are dealing with the conditions for additivity instead of sub-additivity. For two complementary graphs it was proved by Csiszár *et al.* that additivity for every possible probability distribution is equivalent to the perfectness of the involved graph(s). here we give necessary and sufficient conditions of additivity (for every prob. dist.) in full generality. This can be considered as a generalization of the concept of perfect graphs, giving some insight into the kind of graph properties to which graph entropy is sensitive.

Maximum Growth Exponent Equals Minimum Information Rate

Paul Algoet *Information Systems Lab, Durand 141B, Stanford, CA 94305*

We formulate a convex duality principle and prove that the maximum growth exponent, taken over a convex family of functions, is equal to the minimum information rate, taken over the polar family of measures. This minimax property remains valid when sequences are selected from convex families so as to maximize the asymptotic growth exponent. The results are related to the variational characterizations

of relative entropy rate that is due to Donsker and Varadhan, and to the asymptotic optimality principle for log-optimum investment and gambling that was formulated by Algoet and Cover.

Moving Average Processes and Maximum Entropy

Dimitris Nicolas Politis *Department of Statistics, Stanford University, Stanford, CA 94305-4020*

Let X_t, $t \in \mathbf{Z}$ be a wide sense stationary stochastic process with mean $EX_t = 0$ and autocovariance $\gamma(k) = EX_tX_{t+k}$, $k \in \mathbf{Z}$. It is well known [Burg 1967] that the Maximum Entropy such process that satisfies the constraints $\gamma(i) = c_i$, $i = 0, 1, \ldots, p$ is the AR (autoregressive) Gaussian process that satisfies these constraints. Physical or practical considerations might in some cases impose the additional constraint that $\gamma(i) = 0$, $i > q$. Since any time series with $\gamma(i) = 0$, $i > q$, is a MA (moving average) process of order (at most) q, we then face the problem of finding the Maximum Entropy process among the MA(q) processes that satisfy the constraints $\gamma(i) = c_i$, $i = 0, \ldots, p$. The solution to this problem rests upon the relationship between the autocorrelation and inverse autocorrelation function of an AR process that was recently brought to light by Kanto (1987). It is to be noted that in the context of spectral estimation, $q = p$ corresponds to a periodogram-like estimator, while $q = \infty$ leads to Burg's all-pole (AR) estimator. Hence the choice $p < q < \infty$ yields a solution intermediate between the periodogram and Burg's AR estimator.

Some Correlation Properties of and Entropy Calculations in 2-D Lattice Filters

A. Ertuzun and E. Panayirci *Department of Electrical and Electronic Engineering Bogazici Universitesi, Bebek, Istanbul, Turkey, and Faculty of Electrical and Electronics Engineering, Ayazaga Kampusu, Istanbul, Turkey*

The theory of one-dimensional (1-D) lattice filters are well developed and well known. In recent years, two dimensional (2-D) lattice filters are being investigated intensively and their theories are being developed. 2-D lattice filters with quarter plane support have many similarities with the causal 1-D lattice filters. Making use of these similarities, the correlation relations between the forward and the backward prediction errors are derived. However, there are some differences between the 2-D and the 1-D lattice filters due to the differences in the growth of data support. Even though the data support of a 1-D lattice filter grows linearly with the order of the filter, the growth is not linear in a 2-D case since the support grows in both directions. Thus a lattice filter with fixed number of coefficients is not sufficient to model a 2-D autoregressive (AR) data field. The information is lost while modeling a data field with fixed number of reflection coefficients. An expression is derived to express the autocorrelation matrix of the backward prediction errors in terms of the reflection coefficients and the residual error power at each stage. The entropy of the backward prediction errors is calculated in terms of its autocorrelation matrix and it is compared with that of the data field. The entropy lost is calculated quantitatively by computer simulations and the theory is confirmed.

SESSION FA6

CODING THEORY IX

Multilevel Codes with Bounded M-th Order Running Digital Sum

E. Eleftheriou and R. Cideciyan *IBM Research Division, Zürich Research Laboratory, CH-8803 Rüschlikon, Switzerland*

Multilevel sequences with a spectral null of order M, that is, the power spectral density and its first $2M-1$ derivatives vanish, are characterized by finite-state transition diagrams, whose edge labels satisfy bounds on the variation of the M-th order running digital sum. Necessary and sufficient conditions for sequences exhibiting a spectral null or order M are given. For this new class of codes a lower bound on the minimum Euclidean distance at the output of partial response channels with spectral null of order P is obtained. It is shown that the distance bound depends on the sum of the orders of code and channel spectral nulls and can be met with equality provided that $M+P \leq 10$. The case of $M+P > 10$ leads to an unsolved problem in number theory. Examples of quaternary codes on dicode channels and their maxentropic power spectral density are given.

Parallel and Variable Coding/Decoding of MDS-Codes

B. G. Dorsch *Technische Hochschule Darmstadt, Institut für Netzwerk- und Signatheorie, Merckstrasse 25, D-6100 Darmstadt, F.R. Germany*

Usual coding/decoding algorithms for Maximum-Distance-Separable (MDS)-Codes, as Reed-Solomon (RS)- or extended RS-codes, either are not variable in codeparameters (n,k), n = codelength, k = dimension, and/or don't have a parallel structure as required for fast processing. New algorithms with variable and parallel structure based on Newton's interpolation of polynomials are described, suitable for general purpose RS-coders/decoders (over a fixed field $GF(q)$), including simplifications for correction of errors and erasures. The corresponding syndromes can be used for arbitrarily shortened, extended permutated or punctured RS-codes either with the Berlekamp-Massey-Algorithm (BMA) or the Euclidian-Division Algorithm (EDA) to calculate error-location- and error-value-polynomial. Simple relations to the usual Discrete-Fourier-Transform (DFT)-representation of MDS-codes are given. Further simplifications of coding/decoding for codes over the field of real or complex number (for 'coded modulation') are indicated.

A New Table of Constant Weight Codes

A. E. Brouwer, James B. Shearer, N. J. A. Sloane, and Warren D. Smith *Technological University of Eindhoven, 5600 MB Eindhoven, Netherlands; IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598; and AT&T Bell Laboratories, Murray Hill, NJ 07974.*

This paper presents a table of binary constant weight codes of length $n \leq 28$. Explicit constructions are given for almost all the 600 codes in the table; most of these codes are new. We briefly survey the known techniques for constructing constant weight codes, and also give a table of (unrestricted) binary codes of length $n \leq 28$.

The Nonexistence of t–QP Codes, for $t > 2$, and Some New 2–QP Codes

Behnam Kamali and Harold Longbotham *Division of Engineering, University of Texas at San Antonio, San Antonio, TX 78285*

Quasi-perfect (QP) codes are optimum for the binary symmetric channel in the absence of perfect codes. We prove that there are no linear t-QP codes for $t > 2$, where t is the number of errors the code can correct. This proof is based on the properties of parity check matrices for linear block codes. A new

search method for 2-QP codes is described. Several new 2-QP codes, found using this method, are reported.

Construction of Optimal or Nearly Optimal m-Out-Of-n Codes Through Arithmetic Coding

Tenkasi V. Ramabadran *Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa 50011*

A (n,k) block code in which each binary codeword of length n has exactly m ls is referred to as a m-out-of-n code. Such codes are useful, for example, in providing *perfect* error detection capability in asymmetric channels such as optical communication links and laser disks. A m-out-of-n code is *optimal* when, for a given number of information bits k, and m, the number of check bits r $(=n-k)$ is minimum (r_{opt}). The code is *nearly optimal* when $r = r_{opt} + 1$. In this presentation, we describe a new scheme for the construction of m-out-of-n codes based on the arithmetic coding technique. Basically, arithmetic decompression is used to encode a k-bit information word into a n-bit codeword and arithmetic compression is used for the decoding operation. The encoding and decoding algorithms of the scheme involve only simple arithmetic operations performed recursively, thereby facilitating the construction of codes with relatively *long* block sizes. The scheme allows the construction of optimal or nearly optimal m-out-of-n codes for a wide range of block sizes limited only by the arithmetic precision used.

Anticode Construction and Bounds on Maximum Distance

Valdemar Cardoso da Rocha, Jr., and Marcia Mahon Campello de Souza *Department of Electronics and Systems, Communications Research Group - CODEC, Federal University of Pernambuco, Cidade Universitária, 50.741 Recife PE, Brazil*

Two techniques are presented for constructing new anticodes from known anticodes, namely product and interleaving of anticodes. The product of two anticodes with parameters (m_{1,k_1}, δ_1) and (m_{2,k_2}, δ_2) respectively, produces an (m,k,δ) anticode, where $m = m_1 m_2$, $k = k_1 k_2$ and $\delta \leq \min[m_1 \delta_2, m_2 \delta_1]$. It is also shown that an interleaving of degree λ of an (m,k,δ) anticode produces an $(m\lambda, k\lambda, \delta\lambda)$ anticode. Many of the anticodes constructed are either optimum or near optimum in the sense of having the lowest maximum distance δ for given values of m and k. Another topic covered in this paper is relative to bounds on the maximum distance δ of an anticode. A Gilbert-Varsharmov type of upper bound and a tighter form of the Plotkin lower bound on the maximum distance of linear anticodes are presented.

Nonsystematic d-Unidirectional Error Detecting Codes

Eiji Fujiwara and Masayuki Sakura *Dept. of Computer Science, Tokyo Institute of Technology, Ookayama, Meguro-Ku, Tokyo 152, Japan*

This paper presents a simple construction method for the optimal nonsystematic d-unidirectional error detecting (d-UEC) codes. A necessary and sufficient condition for the d-UED codes, C, is stated such that for all $X, Y \in C$ either X and Y are unordered or the Hamming distance between them is at least $d+1$ when one covers the other. First we adopt the weight zero codeword with length n and then adopt weight $d+1$ constant-weight codewords. Next, codewords with weight $2(d+1)$, $3(d+2)$, $\cdots$, are adopted until the weight is less than or equal to the codeword length, n. In this case, it can be easily found that the number of codewords can be varied by adopting the nonzero codewords as the first step in the above construction method. We conclude that adopting weight $|\lfloor n/2 \rfloor|_{d+1}$ codewords as the initial set of codewords in the above method and then adding codewords with weight $d+1$ to the previous ones gives the maximum number of codewords of the d-UED codes, which is equal to the theoretical bounds. ($|X|_y$ means the residue of X divided by y, and $\lfloor x \rfloor$ means the maximum integer no greater than x.)

A New Class of Constructive Asymptotically Good Generalized Concatenated Codes Beyond the Zyablov Bound

Toshihisa Nishijima, Hiroaki Ishii, Hiroshige Inazumi, and Shigeichi Hirasawa *Faculty of Engineering, Kanagawa Institute of Technology, Kanagawa, 243-02 Japan, Sagami Institute of Technology, Kanagawa, 251 Japan, and School of Science and Engineering, Waseda University, Tokyo 169, Japan*

A new class of constructive asymptotically good codes are proposed by using the construction method of the generalized concatenated codes. The outer code of the codes is formed from $J (J \geq 2)$ Reed-Solomon codes with the same code length and the different code rates. The inner code of the codes is the interleaved code constructed by the base code with known weight distribution, proposed by E. J. Weldon, Jr., where the inner code is a nonsystematic code and has $J-1$ subcodes. The lower bounds for the new codes are the best bounds for code rate r, $0.0741 < r < 0.344$, in all the constructive asymptotically good codes previously known to us. Moreover, we obtain an interesting result that those for the new codes lies more above the Zyablov bound for code rate, $0.0838 < r < 0.354$.

SESSION FA7

CONVOLUTIONAL CODES

On Bit-Error Probability for Convolutional Codes

Marat V. Burnashev and David L. Cohn *Institute for Problems in Information Transmission, USSR Academy of Sciences, Ermolovoy str. 19, 101447, Moscow, and Department of Electrical & Computer Engineering, University of Notre Dame, Notre Dame, IN 46556* (40 min.)

A new analytic expression for bit-error probability for convolutional codes is proposed. The expression uses the new concept of *basic metric state* and the associated notion of a *new time scale*. These allow the behavior of a decoder of convolutional codes to be viewed, to a certain extent, like a decoder of block codes. The expression is derived for performance over the binary symmetric memoryless channel by rate $1/n$ codes. It is used to find improved upper bounds on bit-error probability and to analyze its asymptotic behavior.

A Class of Self-Orthogonal Convolutional Codes

Valdemar C. da Rocha, Jr. *Department of Electronics and Systems, Communications Research Group - CODEC, Federal University of Pernambuco, Cidade Universitária 50.741, Recife PE, Brazil*

This paper presents a systematic procedure, based on linear congruences, for the construction of a class of binary self-orthogonal convolutional (n,k,m) codes of minimum distance $d = n - k + 1 = J + 1$, where $k = m + 1 = p$, $1 \leq J \leq p$, and p is a prime number.

Ring Convolutional Codes for Phase Modulation

James L. Massey, Thomas Mittelholzer, Thomas Riedel, and Mark Vollenweider *Inst. for Signal and Info. Proc., Swiss Federal Institute of Technology, 8092 Zurich, Switzerland*

It is argued that "linear codes" over the ring of integers modulo M are the natural linear codes for use with phase modulation. A ring weight and a ring distance can be defined such that the ring distance between two codewords equals the ring weight of their difference and also equals the squared Euclidean distance between the corresponding sequences of modulated signals. Moreover, a phase shift of $360/M$ degrees in a modulated signal corresponds to adding 1 to the associated ring element. A list is given of ring convolutional codes for 4-phase and 8-phase modulation that have generally at least as large free Euclidean distance as the best trellis codes with the same number of states found by applying subset partitioning to ordinary convolution codes but are, in contrast to the latter codes, invariant to a phase shift of $360/M$ degrees.

An appropriate theory of M-ary (n,k) ring convolutional codes is developed. The code is defined as a rank n free module of n-tuples with entries in the ring of M-ary casual rational functions. When M is the e-th power of a prime p, the set of code sequences for an (n,k) ring code coincides with that of an ordinary (ne,ke) convolutional code over $GF(p)$; a linear encoder for the ring code is catastrophic and/or minimal according as its field code counterpart is catastrophic and/or minimal. A simpler test for catastrophity of the ring encoder is given. It is shown that (unlike in the field case) a ring convolutional code can have no non-catastrophic polynomial encoder and can have no minimal polynomial encoder, but (like in the field case) always has a non-catastrophic and minimal systematic encoder.

A Convolutional Decoding Structure for High Data Rate Applications

R. Schweikert and A. J. Vinck *German Aerospace Research Establishment (DLR), Institute for Communication Technology, Oberpfaffenhofen, D-8031 Wessling, Federal Republic of Germany, and Eindhoven University of Technology, Information-Communication Theory, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands*

We describe a concatenated coding scheme suitable for high data rate applications. The structure of the considered scheme basically consists of a set of parallel operating short constraint length $K = 3$ convolutional codes as inner code and a single parity check outer code. Consequently the first step in the decoding is performed by a set of parallel operating 4-states Viterbi-decoders. The outputs of these decoders are used as input for the soft-decision single parity check outer decoding. By using the analogue demodulator outputs the outer decoder is able to correct one erroneous inner Viterbi-decoder output. The embodiment of the calculations to facilitate this operation is shown to be of low additional complexity.

The structure of the resulting decoder is characterized by its high parallelism. The overall hardware to implement the decoder is roughly the same as that of a $K = 7$, 64-states Viterbi-decoder. However, there is a significant gain in decoding speed. Furthermore the error event length distribution is basically determined by the event length distribution of a $K = 3$ Viterbi-decoder.

The derived upper-bound for the decoding bit-error-rate shows that approximately the same coding gain as for the well-known constraint length $K = 7$ convolutional code with regular soft-decision 64-states Viterbi-decoding can be expected.

An Upper Bound on the Error Performance of Convolutional Coding with Nonindependent Rayleigh Fading

François Gagnon and David Haccoun *Department of Electrical Engineering, Ecole Polytechnique de Montréal, Campus de l'Université de Montréal, P.O. Box 6079, Succ. "A", Montreal (Quebéc) H3C 3A7*

A new upper bound is proposed for the evaluation of the error performance of coded systems over noninterleaved or partially interleaved Rayleigh fading channels. The correlation between successive received symbols is exploited to bound error performance. The bound allows useful evaluation of coding gains on realistic communication systems without going into lengthy computer simulations. It may also provide a useful tool for the search of good codes on continuous channels with memory.

Bidirectional Algorithms for the Decoding of Convolutional Codes

David Haccoun and Jean Belzile *Department of Electrical Engineering, Ecole Polytechnique de Montréal, P.O. Box 60-79, station "A", Montréal (Quebec) Canada H3C 3A7*

New bidirectional multiple-path tree searching algorithms for the decoding of convolutional codes are presented. These suboptimal decoding algorithms, suited for long constraint length convolutional codes, use several paths, all of equal length, in a bidirectional breath-first tree searching manner. It is shown that this bidirectional exploration alleviates some of the bit error propagation due to the correct path loss of the usual M-algorithm. Computer simulations with $M = 32$, 64, and 128, show that a coding gain of 1 dB at $P_B = 10^{-5}$ can be achieved over Viterbi decoding using the same number of paths.

Some Easily Analyzable Convolutional Codes

Sam Dolinar, Robert McEliece, Fabrizio Pollara, and Henk van Tilborg *Department of Electrical Engineering and Jet Propulsion Lab, California Institute of Technology, Pasadena, CA 91125, and Department of Math and Computing Science at Eindhoven University of Technology, Eindhoven, The Netherlands*

In this paper we will study a class of convolutional codes which we call *zero-run length (ZRL) convolutional codes*. An $(n, 1)$ convolutional code of constraint length K is said to be ZRL if the encoder's output weight depends only on the input and on the *type* of the state, where the type of a $K - 1$

bit state is defined to be the number of leading 0's in the shift-register contents defining the state. We shall see that ZRL codes include as special cases orthogonal convolutional codes, the recent "superorthogonal codes" of Viterbi and many others. We will find that for any ZRL code, it is possible to compute the free distance by inspection, and to write down the complete transfer function $T(D,IL,L)$, explicitly. In many cases, ZRL codes achieve the largest possible, or nearly the largest possible, free distance for their parameters. Furthermore, the ideas we have developed to analyze these specialized codes may well extend to a much larger class.

AUTHOR INDEX

M

N

O

X

Y

Z